

EPIDEMIOLOGIA & ESTADÍSTICA

S237e Santos, Érica Marques da Silva
Epidemiologia & estatística. / Érica Marques da Silva Santos;
Mariana de Lazzari Gomes (rev.). – Muriaé: FAMINAS, 2024.
117p.

ISBN: 978-65-89983-26-2

1. Epidemiologia. 2. Estatística. I. Santos, Érica Marques da
Silva. II. Gomes, Mariana de Lazzari (rev.). III. Título.

CDD: 614.4

SUMÁRIO

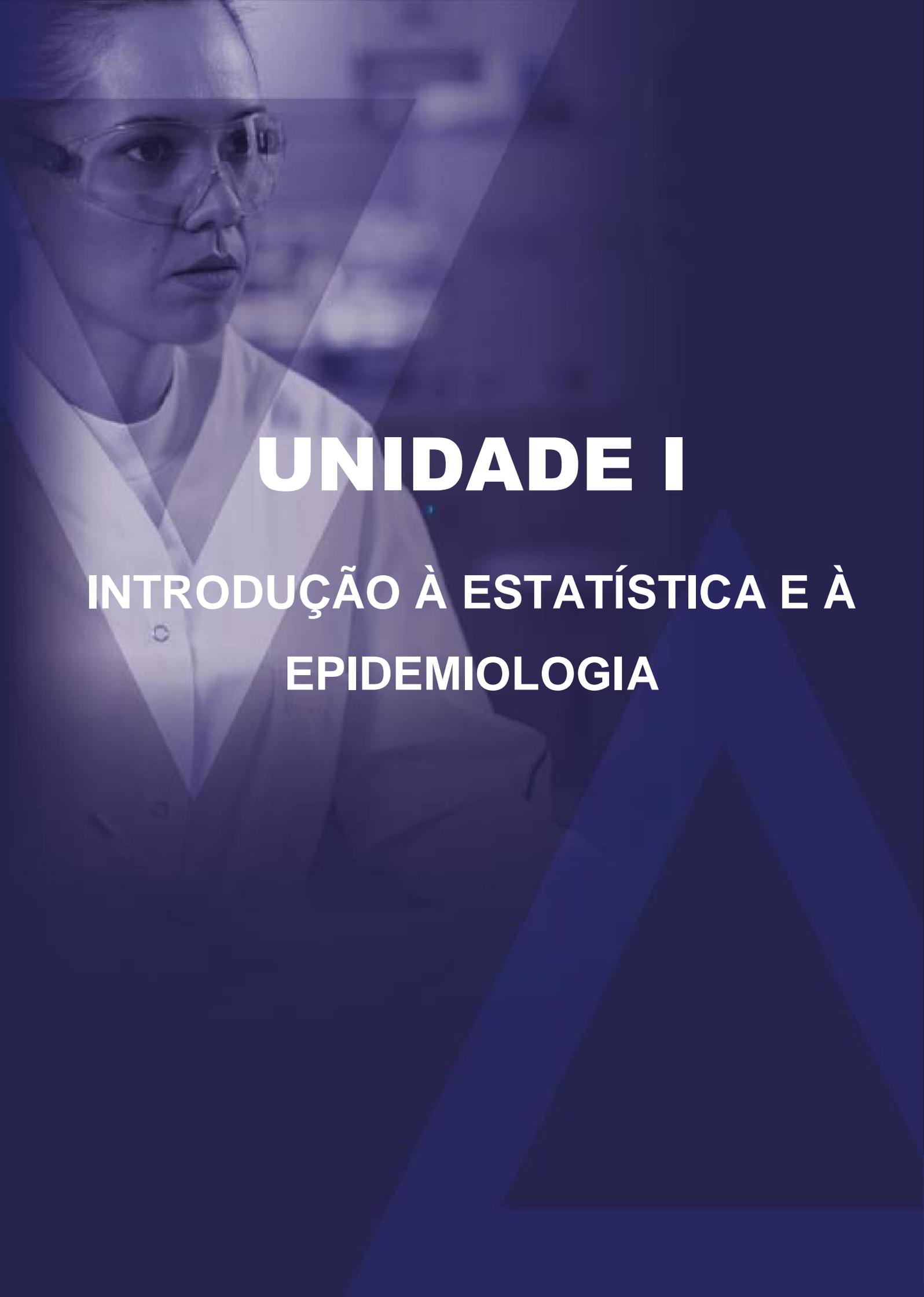
UNIDADE I.....	8
INTRODUÇÃO À ESTATÍSTICA E À EPIDEMIOLOGIA	8
OBJETIVOS.....	11
META DA UNIDADE	11
POPULAÇÃO E AMOSTRA.....	11
VARIÁVEIS ESTATÍSTICAS.....	15
PROPORÇÃO E PROBABILIDADE.....	17
EVENTOS MUTUAMENTE EXCLUSIVOS.....	19
EVENTOS NÃO MUTUAMENTE EXCLUSIVOS.....	20
EVENTOS INDEPENDENTES.....	21
EVENTOS DEPENDENTES	22
INFERÊNCIAS E CONCLUSÕES.....	22
PARA SABER.....	23
EPIDEMIOLOGIA: HISTÓRICO, CONCEITOS E FUNDAMENTOS EPIDEMIOLOGIA; Processo Saúde Doença. História natural da doença	24
BOM, E COMO A EPIDEMIOLOGIA SURTIU?	26
PROCESSO SAÚDE-DOENÇA	27
HISTÓRIA NATURAL DA DOENÇA	28
AS FASES DO MÉTODO ESTATÍSTICO.....	31
RESUMO DA UNIDADE	33
UNIDADE II.....	34
AMOSTRAGEM E ESTUDOS EPIDEMIOLÓGICOS	34
OBJETIVOS.....	35
META DA UNIDADE	35
TIPOS DE AMOSTRAS PROBABILÍSTICAS.....	36
AMOSTRAGEM ALEATÓRIA SIMPLES	37

INDICAÇÃO DE VÍDEO	38
AMOSTRAGEM SISTEMÁTICA.....	38
AMOSTRAGEM ESTRATIFICADA	39
INDICAÇÃO DE VÍDEO	40
AMOSTRAGEM POR CONGLOMERADOS	41
INDICAÇÃO DE VÍDEO	41
PRINCIPAIS TIPOS DE ESTUDOS EPIDEMIOLÓGICOS EXPERIMENTAIS E OBSERVACIONAIS.....	42
ESTUDOS OBSERVACIONAIS.....	42
ESTUDOS DESCRITIVOS.....	43
ESTUDO ECOLÓGICO	43
ESTUDOS ANALÍTICOS	44
ESTUDO DE COORTES (GRUPOS).....	45
ESTUDOS EXPERIMENTAIS.....	46
ENSAIO CLÍNICO ALEATÓRIO.....	47
ENSAIOS DE CAMPO.....	47
ENSAIOS COMUNITÁRIOS	48
RESUMO DA UNIDADE	49
UNIDADE III.....	51
ESTATÍSTICAS DESCRITIVAS	51
OBJETIVOS.....	52
METAS	52
TABELAS DE FREQUÊNCIAS E GRÁFICOS.....	53
APRESENTAÇÃO DE DADOS CONTÍNUOS EM TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS	53
APRESENTAÇÃO DE DADOS DISCRETOS EM TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS.....	56
APRESENTAÇÃO DE DADOS CATEGORIZADOS EM TABELA DE DISTRIBUIÇÃO DE	

FREQUÊNCIAS	57
APRESENTAÇÃO DE DADOS QUALITATIVOS	58
APRESENTAÇÃO DE DADOS QUALITATIVOS	60
LEITURA COMPLEMENTAR	62
MEDIDAS DE POSIÇÃO OU TENDÊNCIA CENTRAL	62
MÉDIA	63
INDICAÇÃO DE VÍDEO	63
INDICAÇÃO DE VÍDEO	64
INDICAÇÃO DE VÍDEO	64
MEDIANA	65
INDICAÇÃO DE VÍDEO	65
INDICAÇÃO DE VÍDEO	66
MODA	66
INDICAÇÃO DE VÍDEO	67
INDICAÇÃO DE VÍDEO	67
MEDIDAS DE DISPERSÃO/VARIABILIDADE	68
DESVIO PADRÃO	68
COEFICIENTE DE VARIAÇÃO	69
MEDIDAS DE POSIÇÃO RELATIVA E DIAGRAMA DE CAIXA	69
ESCORES Z	70
INDICAÇÃO DE VÍDEO	70
PERCENTIL	70
QUARTIS	71
INDICAÇÃO DE VÍDEO	72
DIAGRAMA DE CAIXA	72

RESUMO DA UNIDADE	74
UNIDADE IV	75
MEDIDAS EPIDEMIOLÓGICAS BÁSICAS	75
OBJETIVOS.....	76
META DA UNIDADE	76
MEDIDAS DE MORBIDADE	77
PREVALÊNCIA.....	78
INCIDÊNCIA	79
MEDIDAS DE MORTALIDADE	80
INDICADORES DE MORTALIDADE.....	80
TAXA DE MORTALIDADE GERAL	81
TAXA DE MORTALIDADE AJUSTADA.....	81
MÉTODO DE PADRONIZAÇÃO DIRETA	81
MÉTODO DE PADRONIZAÇÃO INDIRETA.....	82
TAXA DE MORTALIDADE INFANTIL	83
TAXA DE LETALIDADE	84
RESUMO DA UNIDADE	86
UNIDADE V	88
INTRODUÇÃO A OS TESTES DE HIPÓTESES	88
OBJETIVOS.....	89
META DA UNIDADE	89
CONCEITOS BÁSICOS DO TESTE DE HIPÓTESE	90
TIPOS DE TESTE DE HIPÓTESE: BILATERAL, UNILATERAL À ESQUERDA, UNILATERAL À DIREITA.....	92
INDICAÇÃO DE VÍDEO	93
TIPOS DE ERROS.....	93
INDICAÇÃO DE VÍDEO	94

NÍVEL DE SIGNIFICÂNCIA	94
NÍVEIS CLÁSSICOS DE SIGNIFICÂNCIA	94
DEFINIÇÃO E INTERPRETAÇÃO DE P-VALOR.....	96
INDICAÇÃO DE VÍDEO	96
ERRO TIPO 2 E PODER DO TESTE.....	97
UNIDADE VI	99
MEDIDAS DE EFEITO	99
OBJETIVOS.....	100
METAS DA UNIDADE.....	100
RAZÃO DE RISCO (RR)	100
RR NOS PRINCIPAIS TIPOS DE ESTUDO	101
PROBABILIDADE E CHANCE	102
Cálculo do RR em Estudos de Coorte	102
RAZÃO DAS CHANCES OU <i>ODDS RATIO</i>	104
PARA SABER	107
RAZÃO DE PREVALÊNCIAS.....	107
OUTRA DEFINIÇÃO PARA RAZÕES DE PREVALÊNCIAS.....	108
RESUMO DA UNIDADE.....	109
REFERÊNCIAS	110



UNIDADE I

**INTRODUÇÃO À ESTATÍSTICA E À
EPIDEMIOLOGIA**

APRESENTAÇÃO DA DISCIPLINA

Bem-vindos à disciplina de Estatística e Epidemiologia, aqui você irá adquirir uma base sólida que tornará você capaz de identificar problemas locais e direcionar as intervenções, e reconhecer que isto representa um dos objetivos mais importantes da pesquisa em saúde. A estatística auxilia na produção do novo modelo de profissional, exigido pela área da saúde moderna. Entre outros pressupostos, devem se formar profissionais capacitados para a leitura crítica dos fenômenos que ocorrem na população relacionados ao processo saúde-doença, bem como selecionar e adotar com segurança novas tecnologias divulgadas rotineiramente em seu meio, tais como a formulação de novos medicamentos, novas técnicas cirúrgicas, identificação de grupos ou de fatores de risco para uma doença e alterações nos parâmetros de diagnósticos. A epidemiologia buscar comparar e quantificar o comportamento dos fenômenos de saúde/doença, e seus fatores condicionantes e determinantes, nas populações. Dados estatísticos auxiliam no processo de tomada de decisão em problemas epidemiológicos, sendo assim ao concluir esta disciplina você será capaz de tomar decisões de forma assertiva.

Para que seu estudo seja leve e proveitoso esta disciplina foi organizada em 6 unidades, com temas e subtemas que, por sua vez, são subdivididos em seções (tópicos), atendendo aos objetivos do processo de ensino-aprendizagem.

Esperamos que, até o final da disciplina vocês possam:

Correlacionar os conceitos estatísticos e epidemiológicos básicos para interpretação de dados de pesquisas na área da saúde.

Aplicar os conhecimentos no desenvolvimento, implementação e análise de pesquisas e estudos científicos.

Mas antes de iniciarmos nossos estudos, convido você a pensar nas seguintes situações:

Você deseja investigar a relação entre consumo de cigarros e câncer de pulmão. Esta hipótese de que o consumo de cigarros tenha associação com câncer de pulmão pode ser verificada, determinando-se o risco para tabagistas, baseando-se em indivíduos não fumantes, tendo-se em conta que os não tabagistas também podem contrair a doença. Conceituar o que seja risco é fundamental. Portanto, considera-se como risco a probabilidade de que pessoas que estão sem a doença, mas expostas a certos fatores, possam adquiri-la.

Na verificação da ocorrência de doentes na população, variável chave da investigação epidemiológica, as medidas de incidência, prevalência e outros coeficientes servem como indicadores de risco.

Agora vamos lá, na situação relatada acima, há termos que provavelmente você já tenha ou não ouvido falar, como: população, variável, medidas de incidência, prevalência. Estes são alguns conceitos que iremos explorar nesta disciplina.

Bom, acho que já deu pra ter uma ideia da importância desta disciplina para seu crescimento profissional e intelectual, então agora convido você e embarcar comigo nesta aventura de desvendar os mistérios e desafios da estatística e epidemiologia.



OBJETIVOS

- ✓ Introduzir os conceitos e definições básicas da estatística e da epidemiologia.



META DA UNIDADE

Na área da saúde humana e animal, a estatística, especificamente, pode ser entendida em dois ambientes. O primeiro se refere ao levantamento de informações, como registro de doenças, surtos, endemias, epidemias e de registros de qualidade de vida, como condições de alimentação, sanitárias, habitacionais, de prevenção de doenças, educação etc. Esse ambiente é conhecido como ambiente macro e tem a ver, fundamentalmente, com a identificação, a planificação e a execução de ações de Saúde Pública.

O segundo ambiente se refere à elaboração de experiências e pesquisa científica, tais como testes de vacinas, avaliação de terapêuticas e tratamentos, testes de medicamentos etc. Denomina-se esse ambiente de ambiente micro e tem a ver, naturalmente, com a pesquisa laboratorial e científica.

Nesta unidade abordaremos os conceitos iniciais que constituem a base para a compreensão da estatística e da epidemiologia de forma prática e aplicada.

POPULAÇÃO E AMOSTRA

Provavelmente, quando você lê ou ouve falar na palavra população, você logo associar o termo população a um conjunto de pessoas, não é verdade? Em estatística, o sentido da palavra se torna mais amplo.

Entende-se por população a totalidade dos elementos ou de um atributo dos elementos referentes a um conjunto determinado. Vejamos exemplos para compreender melhor:

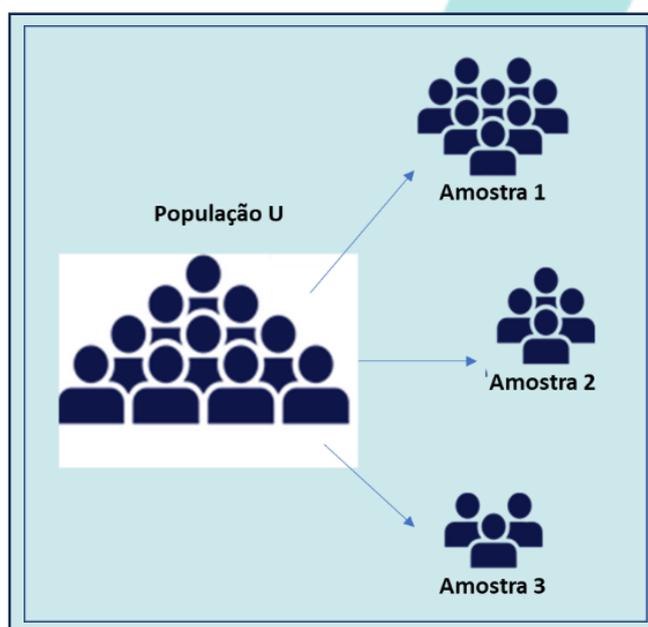
- ✓ População de Muriaé, cujos elementos são as pessoas que residem nessa cidade.
- ✓ População de pacientes internados na FCV, que tem como elementos as pessoas internadas na Fundação Cristiano Varella de Muriaé (FCV).
- ✓ População dos cavalos Mangalarga Marchador machos do Haras Riacho Grande, que tem como elementos cavalos machos do Haras Riacho Grande...
- ✓ População de seringas descartáveis do Posto de Saúde do Bairro do Cruzeiro, cujos elementos são seringas.

Em primeiro lugar, nota-se que a população estatística não é necessariamente um conjunto de pessoas. Por outro lado, percebe-se que a população pode ser enumerável finita, como o caso do primeiro, terceiro e quarto exemplos citados. A população pode ser também infinita, ou impossível (difícil) de enumerar, como o caso do segundo exemplo, no qual, por falta de maiores informações a respeito do conjunto, torna-se difícil a sua contagem.

A dificuldade em enumerar ou tratar conjuntos completos de dados faz com que se trabalhe com partes do conjunto original, tidas como representantes do conjunto. Convencionou-se denominar essas partes como amostras. Desse modo, uma amostra é uma parte tomada da população, ou um conjunto de elementos da população selecionado segundo algum critério, como será abordado aqui.

Considerando a população como um conjunto universal U , uma amostra pode ser

FIGURA 1 - EXTRAÇÃO DE AMOSTRAS DE UMA POPULAÇÃO



FONTE 1: ELABORAÇÃO PRÓPRIA

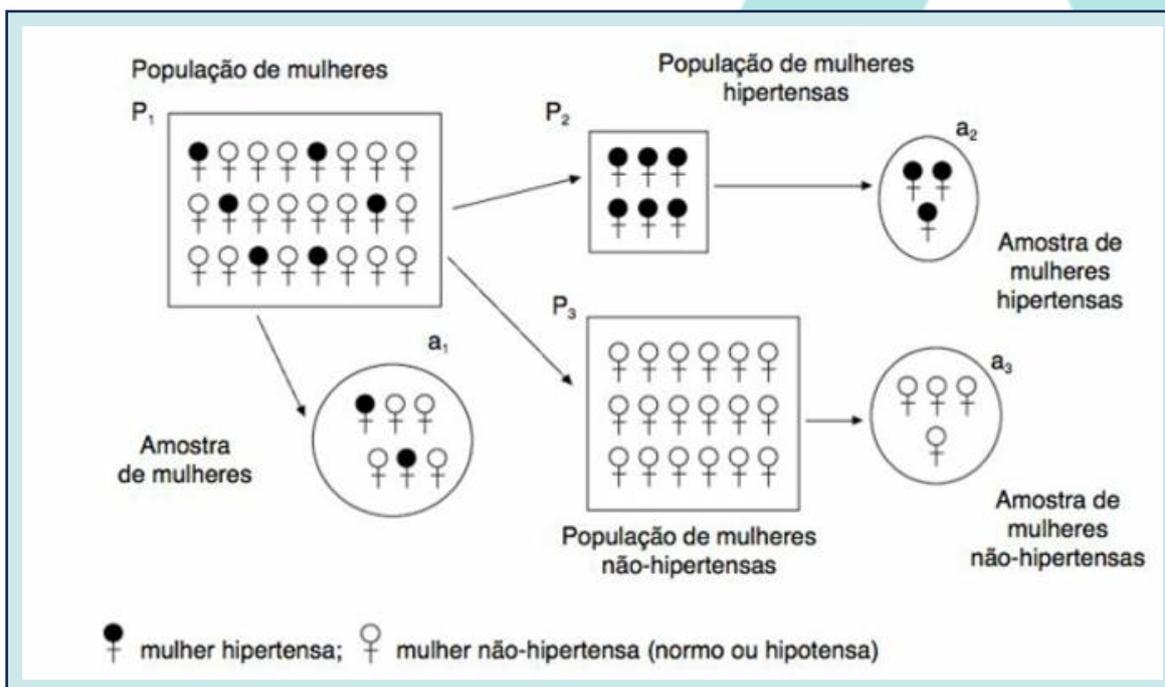
considerada como um subconjunto de U. É claro que várias amostras também podem ser extraídas de uma mesma população, como sugere a FIG. 1.

Em estatística, a definição correta da população-alvo de um estudo é uma questão de grande importância para a pesquisa. Como foi possível observar nos exemplos anteriores, informações acerca da abrangência espacial (geográfica) e temporal servem para definir a população com maior precisão. Como regra, quando é acrescentada a um conjunto de dados uma característica adicional, está-se definindo uma nova população, mais específica (ou menos geral) que a anterior. É importante notar que essa nova população é outra população.

Por exemplo, imagine que estão sendo estudados os efeitos colaterais de determinado medicamento oral. O objeto de estudo pode ser definido, da forma mais abrangente possível, como a população feminina. Assim, a população seriam todas as mulheres. Contudo, não é sempre fácil ou possível relacionar os efeitos da droga com toda a população. Pode ser, por exemplo, que determinado subgrupo feminino seja mais sensível a determinado efeito.

Suponha que esse subgrupo seja o das mulheres hipertensas. Observa-se que, ao se acrescentar uma característica ao grupo “mulheres”, tornando-o o grupo “mulheres hipertensas”, foi definida uma outra população, diferente da anterior, mesmo que os indivíduos do grupo original estejam necessariamente todos no novo grupo. Não se deve pensar que o novo grupo é uma amostra do primeiro simplesmente porque é um subconjunto dele. A razão disso é que o grupo “mulheres hipertensas” possui todas as mulheres que têm pressão arterial

FIGURA 2 - DEFINIÇÃO DE POPULAÇÃO E AMOSTRA



FONTE 2 - ARANGO, 2009, P. 4.

elevada. Portanto, é uma população. Para compreender melhor, observe a FIG. 2, que representa, de forma esquemática, a situação relatada (ARANGO, 2009).

CURIOSIDADE

Você pode encontrar uma aplicação do que aprendemos em uma bula de remédio. Quer conferir? Leia a bula de um remédio qualquer e verifique as indicações, as contraindicações ou os efeitos adversos (efeitos colaterais ou efeitos secundários). Observe como as informações estão direcionadas a grupos específicos. Perceba que todas essas informações foram testadas em amostras de cada um dos subgrupos (populações). Verifique se existem restrições de caráter geral, ou seja, indicações sem referência a um grupo específico.

Agora que você já sabe a diferença entre população e amostra, é importante também conhecer e compreender os significados dos termos estatística e parâmetro. Esses termos são usados para distinguir entre casos nos quais temos dados de uma amostra e casos em que temos dados de uma população inteira.

- Um parâmetro é uma medida numérica que descreve alguma característica de uma população.
- Uma estatística é uma medida numérica que descreve alguma característica de uma amostra.

Dessa forma, podemos dizer que a estatística tem duas definições, aplicadas em contextos diferentes:

- Definição 1. Estatística é uma medida numérica que descreve característica de amostras.
- Definição 2. Estatística é a ciência de planejamento de estudos e experimentos; obtenção de dados; organização, resumo, apresentação, análise e interpretação desses dados; e, então, a extração de conclusões com base nisso (TRIOLA, 2017).

Podemos determinar qual das duas definições se aplica, considerando o contexto no qual o termo estatística é usado. Veja:

EXEMPLO

“Em uma pesquisa da Harris Poll, 2.320 adultos nos Estados Unidos foram inquiridos sobre piercings corporais e 5% dos respondentes disseram ter um piercing, mas não na face. Com base nos últimos dados disponíveis no momento da escrita deste livro, há 241.472.385 adultos nos Estados Unidos. Os resultados da pesquisa são uma amostra extraída da população de todos os adultos.” (Triola, 2017)

1.Parâmetro: O tamanho da população de 241.472.385 é um parâmetro, porque se baseia na população inteira de todos os adultos nos Estados Unidos.

2.Estatística: O tamanho amostral de 2.320 adultos pesquisados é uma estatística, porque se baseia em uma amostra, não na população inteira de todos os adultos nos Estados Unidos. O valor de 5% é outra estatística, porque também se baseia na amostra, não em toda a população.

VARIÁVEIS ESTATÍSTICAS

Uma variável pode ser definida como uma característica de um conjunto de elementos, desde que essa característica apresente mais de uma classificação possível. As variáveis, de acordo com a forma como são medidas, admitem a seguinte classificação:

- Variáveis quantitativas ou numéricas: são avaliadas de forma numérica.
- Variáveis qualitativas ou literais: são expressas pôr a) palavras (atributos), como por exemplo, tipo sanguíneo, ou b) ordens, como 1º, 2º etc.

Por sua vez, as variáveis quantitativas se dividem em:

- ✓ **Discretas:** admitem somente números inteiros (conjunto dos números Naturais). Exemplos: seriam as variáveis resultantes de contagens, como o número de batimentos cardíacos, a frequência respiratória ou a concentração de plaquetas no sangue.
- ✓ **Contínuas:** admitem números fracionários (conjunto dos números Reais). Estão relacionadas com mensurações. Exemplos: tempo de coagulação, perímetro cefálico e peso do fígado.

Já as variáveis qualitativas se dividem em:

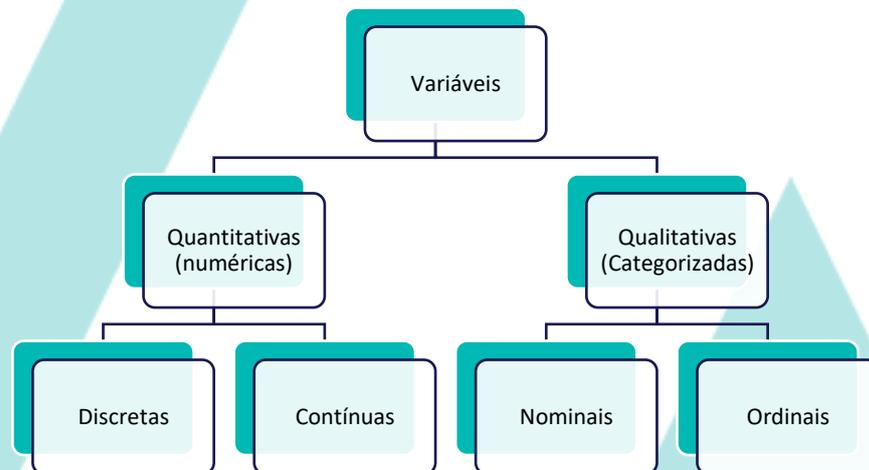
- ✓ **Nominais:** quando podem ser divididas em grupos ou categorias, sem associação numérica. Exemplos: tipo sanguíneo, nacionalidade, raça, espécie, tipo de lesão,

sintomas.

- ✓ **Ordinais:** quando os dados são distribuídos em grupos que têm ordenação. Exemplos: escolaridade, satisfação do cliente, estágio de uma doença, grau de dor.

O esquema apresentado na FIG. 3 mostra como se classificam as variáveis.

FIGURA 3 - CLASSIFICAÇÃO DAS VARIÁVEIS ESTATÍSTICAS



FONTE 3: ELABORAÇÃO PRÓPRIA

Além de conhecer os tipos de variáveis e como elas se classificam, é importante entender a diferença entre variável e dado. Para isso, vamos analisar a seguinte situação: *Carlos Alberto de Lima foi aprovado em primeiro lugar no concurso para os Correios, ele tem 19 anos e deve trabalhar 6 horas por dia. Vejamos:*

Nome: Carlos Alberto Lima é variável nominal;

Lugar de aprovação: 1º, é variável ordinal;

Idade: 19 anos é variável discreta;

Carga horária: 6 horas é variável contínua.

Sendo assim, tivemos quatro variáveis que descreveram o candidato, e os valores e características observados para essas **variáveis** constituem os **dados** sobre o candidato aprovado. Formalmente, podemos dizer que:

- Variável é uma condição ou característica que descreve uma pessoa, um animal, um lugar, um objeto, uma ideia e que assume valores diferentes (é variável) em

diferentes unidades.

- Dados são os valores observados das variáveis.

PROPORÇÃO E PROBABILIDADE

A origem dos estudos de probabilidade se deu através dos jogos de azar, pois todos queriam entender a regra ou lei desses jogos para que fazer fortunas em mesas de jogos e cassinos. Após muitos estudos e incansáveis tentativas, chegou-se à conclusão de que não é possível prever o resultado de um evento aleatório. Podemos apenas conhecer quais as chances de sucesso, mas quando acontecerá, não é possível. A estatística, por sua vez, utiliza dos conceitos de probabilidade para garantir que os resultados sejam confiáveis e precisos (VIEIRA, 2018).

Agora, vamos aprender um pouco sobre probabilidade. Venha comigo!

Ao estudar a probabilidade, temos o espaço amostral e o evento, que podem ser definidos como:

- ✓ **Espaço amostral:** lista com todos os resultados possíveis de um procedimento.
- ✓ **Evento:** resultado ou um conjunto de resultados com determinada característica.

Veja o seguinte exemplo:

Ao jogar um dado



Os resultados possíveis são



Imagine que se deseja obter a face 4 voltada para cima.



A face desejada constitui o evento, ou seja, um resultado desejado.

Agora que já vimos o que é espaço amostral e evento, podemos definir, formalmente, probabilidade. Então, probabilidade de ocorrer um evento com determinada característica é dada pelo quociente entre o número de eventos favoráveis e o número de eventos possíveis.

$$\text{Probabilidade} = \frac{n^{\circ} \text{ de eventos favoráveis}}{n^{\circ} \text{ de eventos possíveis}}$$

Os estatísticos preferem expressar valores de probabilidade por números entre 0 e 1 porque, em cálculos mais avançados, isso é necessário. Na prática, porém, é comum fornecer probabilidades em porcentagens. Se você quiser expressar probabilidade em porcentagem, basta multiplicar o valor dado pela definição por 100 e acrescentar o símbolo de porcentagem (%) ao resultado (TRIOLA, 2017).

EXEMPLO

A probabilidade de ocorrer a face 4 quando se joga um dado equilibrado é dada pelo quociente,

$$P(4) = \frac{1}{6} = 0,1667 \text{ ou } 16,67\%$$

Eventos certos têm probabilidade 1, ou 100%. Eventos impossíveis têm probabilidade zero. Por exemplo, a probabilidade de que qualquer um de nós venha morrer um dia é 1 ou 100%, pois é evento certo. A probabilidade de que qualquer um de nós seja imortal é 0 ou, o que é o mesmo, impossível.

A definição dada aqui não permite responder perguntas, tais como qual é a probabilidade de ser aprovado em um concurso? Qual é a probabilidade de chover amanhã? Qual é a probabilidade de uma pessoa chegar aos 100 anos? Não se pode obter a probabilidade por conjecturas. É aí que entra a frequência relativa.

A **frequência relativa** de determinado evento é dada pelo cociente entre o número de vezes que o evento ocorreu (número de ocorrências) e o número de eventos observados (número de observações)

$$\text{Frequência relativa} = \frac{n^{\circ} \text{ de ocorrências}}{n^{\circ} \text{ de observações}}$$

As frequências relativas variam entre 0 e 1, mas são mais bem entendidas quando expressas em porcentagem.

EVENTOS MUTUAMENTE EXCLUSIVOS

Dois eventos A e B são mutuamente exclusivos se não puderem ocorrer a o mesmo tempo. A ocorrência de um deles exclui (impede) a ocorrência do outro.

EXEMPLO

Ao lançar um dado, só pode ocorrer uma das faces. Assim, se ocorreu a face “3”, ficou excluída a possibilidade de ter ocorrido qualquer outra face.

Importante!!! Se A e B são eventos mutuamente exclusivos, a probabilidade de ocorrer A ou B é igual à soma das probabilidades de ocorrer cada um deles.

Escreve-se: $P(A \cup B) = P(A) + P(B)$

EXEMPLO

Quando você joga um dado, só pode ocorrer uma das faces. Qual é a probabilidade de, em um lançamento, ocorrer 1 ou 6? Usando a regra da soma 1, você calcula a probabilidade de ocorrer 1 e a probabilidade de ocorrer 6. Depois, soma essas probabilidades.

$$P(1) = \frac{1}{6}$$

$$P(6) = \frac{1}{6}$$

$$P(1 \cup 6) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

EVENTOS NÃO MUTUAMENTE EXCLUSIVOS

Dois eventos A e B são não mutuamente exclusivos se houver pelo menos um resultado em comum.

Se A e B são dois eventos não mutuamente exclusivos, há uma sobreposição, isto é, existe pelo menos um resultado de A que também é resultado de B.

EXEMPLO

Quando você joga um dado, só pode ocorrer uma das faces. Mas pense no evento “número ímpar” e no evento “número maior do que 4”. Esses dois eventos têm um resultado em comum: é o número 5, que tanto pertence ao evento “número ímpar” como ao evento “número maior do que 4”.

Solução: Os resultados possíveis quando se lança um dado são apresentados em seguida: “números ímpares” estão circundados por uma elipse e “números maiores do que quatro” por um quadrado. Se você contar o número de resultados que correspondem ao evento “número ímpar” e o número de resultados que correspondem ao evento “número maior do que 4”, terá contado 5 duas vezes.



Importante!!! Regra da Soma 2 - Se A e B são eventos não mutuamente exclusivos, a probabilidade de ocorrer A ou B é dada pela probabilidade de A, mais a probabilidade de B, menos a probabilidade de A e B (contada duas vezes).

Escreve-se: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

EXEMPLO

Quando você joga um dado, só pode ocorrer uma das faces. Qual é a probabilidade de, em um lançamento, ocorrer “número ímpar” ou ocorrer “número maior do que 4”? Usando a regra 2 da soma, você calcula a probabilidade de ocorrer “número ímpar”, a probabilidade de ocorrer “número maior do que 4” e a probabilidade de ocorrer “número ímpar maior do que 4”. Depois, aplica a regra da soma 2:

$$\begin{aligned}P(\text{número ímpar}) &= \frac{3}{6} = \frac{1}{2} \\P(\text{número maior que 4}) &= \frac{2}{6} = \frac{1}{3} \\P(\text{número ímpar} \cup \text{número maior que 4}) &= \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6} = \frac{2}{3}\end{aligned}$$

EVENTOS INDEPENDENTES

Dois eventos, A e B, são independentes se a ocorrência de um deles (A ou B) não tiver efeito sobre a ocorrência do outro (B ou A). Na vida real, encontramos muitos exemplos de eventos independentes. Assim, “chover hoje” e “ser feriado amanhã” são eventos independentes porque o fato de “chover hoje” não muda a possibilidade de “ser feriado amanhã”, nem o fato de “ser feriado amanhã” muda a possibilidade de “chover hoje”. Na área de saúde, existem vários exemplos de eventos independentes, como o fato de o estado civil do cidadão não modificar a probabilidade de ele ser calvo.

Importante!!! Regra 1 da multiplicação: Se A e B são eventos independentes, a probabilidade de ocorrer A e B é dada pela probabilidade de ocorrer A, multiplicada pela probabilidade de ocorrer B.

Escreve-se: $P(A \cap B) = P(A) \times P(B)$

EXEMPLO

Você lança dois dados ao mesmo tempo: um é vermelho e outro é amarelo. Para obter a probabilidade de ocorrer a face 3 no dado amarelo e a face 5 no dado vermelho, aplique a regra 1 da multiplicação: calcule a probabilidade de ocorrer a face 3 no dado amarelo e a face 5 no dado vermelho. Depois, multiplique essas probabilidades.

$$P(3 \text{ no dado amarelo}) = \frac{1}{6}$$

$$P(5 \text{ no dado vermelho}) = \frac{1}{6}$$

$$P(3 \text{ no dado amarelo e } 5 \text{ no dado vermelho}) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

EVENTOS DEPENDENTES

Dois eventos, A e B, são dependentes se a ocorrência do evento A modificar a probabilidade de ocorrência do evento B.

Na vida real é comum nos depararmos com exemplos de eventos dependentes, ou seja, eventos que, embora nem sempre sejam a “causa” de outros, aumentam a probabilidade desses outros eventos acontecerem. Por exemplo, o hábito de fumar aumenta a probabilidade de a pessoa ter câncer de pulmão; o motorista alcoolizado tem maior probabilidade de provocar acidente de trânsito.

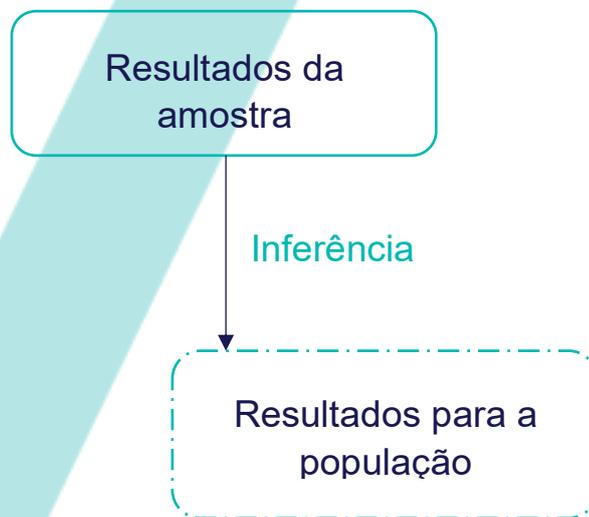
INFERÊNCIAS E CONCLUSÕES

Inferência significa tirar conclusões a partir de dados. A inferência estatística utiliza o método estatístico em dados amostrais e tirar conclusões sobre a população de interesse, descrevendo-a ou testando hipóteses (ZANETTA, s/n). Assim, podemos dizer que:

A inferência estatística é um processo de inferir características de uma população por meio da observação e estudo de uma amostra.

Assim, temos:

FIGURA 4 - ESQUEMA DE INFERÊNCIA ESTATÍSTICA



FONTE 4: ELABORAÇÃO PRÓPRIA,2023

A inferência estatística, utiliza-se do raciocínio indutivo, ou seja, o raciocínio é feito do particular para o geral (amostra para população). A inferência estatística, a partir de informação obtida em uma amostra da população, procura inferir os resultados para a população. Veja que o raciocínio indutivo tem por objetivo ampliar o conhecimento, se as conclusões feitas para os resultados da amostra estiverem corretas.



PARA SABER

Quando é possível obter informação sobre a população inteira, não há necessidade de fazer inferência. Por exemplo, o censo da população brasileira é feito a cada década e todas as pessoas são contadas pelos recenseadores. É possível verificar que, entre os idosos, isto é, aqueles com idade superior a 60 anos, há predomínio de mulheres. A comparação entre a frequência de homens e de mulheres nessa faixa etária é feita diretamente com os resultados obtidos no recenseamento.

EPIDEMIOLOGIA: HISTÓRICO, CONCEITOS E FUNDAMENTOS

EPIDEMIOLOGIA; Processo Saúde Doença. História natural da doença

Você deve estar se perguntando: afinal de contas, o que é EPIDEMIOLOGIA?

Epidemiologia: Histórico, conceitos e Fundamentos em Epidemiologia

A epidemiologia pode ser definida como o ramo da ciência dedicado a estudar “tudo sobre a população”. A origem da palavra é proveniente do grego – **epi (sobre) + demos (população) + logia (estudo)**. De acordo a Organização Mundial da Saúde (OMS), a epidemiologia é o estudo da distribuição e dos determinantes de estados ou eventos relacionados à saúde (incluindo doença) e a aplicação desse estudo ao controle de doenças e outros problemas de saúde (ROUQUAYROL,2017).

Dessa forma, fica claro que os epidemiologistas estão preocupados não somente com a incapacidade, doença ou morte, mas também, com a melhoria dos indicadores de saúde e com maneiras de promover saúde. Assim, por meio da análise da distribuição e dos fatores determinantes das doenças, dos danos à saúde e eventos associados àsaúde coletiva, a epidemiologia visa propor medidas específicas de prevenção, controle ou erradicação de doenças e fornecer indicadores que sirvam de suporte ao planejamento, administração e avaliação das ações de saúde. O alvo de um estudo epidemiológico é sempre uma população humana, que pode ser definida em termos geográficos ou outro qualquer (ROUQUAYROL,2017).

Vários métodos podem ser utilizados para realizar investigações epidemiológicas: **a vigilância e os estudos descritivos** podem ser utilizados para estudar a distribuição; já os **estudos analíticos** são usados para estudar os determinantes. Dessa forma, chegamos à primeira divisão da epidemiologia: descritiva e analítica.

- ✓ **Epidemiologia descritiva:** tem como base a tríade o **tempo**, o **lugar** e as **pessoas**.

Ela se preocupa em realizar descrições gerais relativas à relação de determinada doença com características básicas, como idade, sexo, etnia, ocupação, classe social e localização geográfica, dentro de determinado período. Ela sempre é observacional, nunca experimental.

Um exemplo bastante atual é o levantamento do número de recém-nascidos que foram afetados no Brasil pela epidemia do vírus Zika ocorrida no ano de 2015. Como

podemos visualizar, é feita uma observação da distribuição de determinada doença, mas, em nenhum momento nesse exemplo, cogitou-se a intervenção por meio de uma experimentação científica. Somente é realizada uma coleta de dados. Embora ela apenas descreva os dados, a epidemiologia *descritiva é um antecedente necessário da epidemiologia analítica*(ROUQUAYROL, 2017).

- ✓ **Epidemiologia analítica:** tem como base a tríade **hospedeiro, ambiente e agente**. Por meio do estudo epidemiológico analítico, subconjuntos de uma população definida podem ser identificados, classificando quem esteve, está ou estará *exposto* ou *não exposto* - ou exposto em diferentes graus - a um ou mais fatores que acarretarão no desfecho final – seja ele saúde ou doença.

Para o mesmo exemplo de Zika vírus, mas agora por meio de estudos, verificou-se que o vírus (agente) é transmitido pelo mosquito Aedes aegypti, que se localiza em regiões mais quentes, por isso afetando mais determinadas regiões do país que outras. Além disso, os hospedeiros vertebrados do vírus são humanos e macacos tendo como consequência a ocorrência de microcefalia congênita quando transmitido a gestantes e seus fetos (ROUQUAYROL, 2017).

E aí? Conseguiu perceber a diferença em ambos os casos?

No caso da epidemiologia descritiva, há somente a descrição dos casos, já a epidemiologia analítica busca, por meio de experimentação, explicar os casos como, por exemplo, de que modo a doença é transmitida.

Resumidamente, os objetivos principais dessa ciência são:

Descrever a distribuição e a magnitude dos problemas de saúde nas populações humanas;

Identificar e entender o agente causal e fatores relacionados aos agravos à saúde;

Identificar e explicar os padrões de distribuição geográfica das doenças; estabelecer metas e estratégias de controle;

Estabelecer medidas preventivas;

Auxiliar no planejamento e desenvolvimento de serviços de saúde ao elencar as prioridades.

Importante!!! População: inclui indivíduos com características específicas às quais se deseja estudar. Distribuição: refere-se à análise quanto ao tempo, pessoas, lugares e grupos de indivíduos afetados. Determinantes: inclui fatores que afetam o estado de saúde, dentre os quais os fatores biológicos, químicos, físicos, sociais, culturais, econômicos, genéticos e comportamentais

BOM, E COMO A EPIDEMIOLOGIA SURTIU?

A epidemiologia teve origem na ideia de que fatores ambientais podem influenciar a ocorrência das doenças. Ela se originou das observações de Hipócrates, pai da medicina grega, feitas há mais de 2000 anos, de que fatores ambientais influenciam a ocorrência de doenças.

No entanto, a quantificação dos problemas de saúde se inicia somente no século XVII, com John Graunt, considerado o pai da estatística e da demografia. Por meio de tabulação de dados sobre a mortalidade em Londres, ele conseguiu reunir dados de óbitos por sexo, região, percentuais de mortes em relação aos nascimentos e à população total.

Passados alguns anos, a Revolução Industrial, no século XIX, fez com que ocorresse muitos deslocamentos populacionais das zonas rurais para as áreas urbanas. Esse movimento populacional, que ocorreu de forma desestruturada, contribuiu para a ocorrência de epidemias de cólera, febre amarela, febre tifoide, dentre outras. Foi nesse contexto que John Snow, considerado o pai da epidemiologia, também em Londres, estabeleceu a relação entre o contágio por cólera e o consumo de água contaminada.

Ainda no século XIX, as descobertas de Louis Pasteur sobre microbiologia levam a profundas mudanças na epidemiologia. Ao identificar e isolar numerosas bactérias, Pasteur conseguiu elaborar meios de eliminar as bactérias patogênicas, assim como levou ao desenvolvimento do uso de bactérias em processos industriais e farmacológicos. Robert Koch, dando sequência às pesquisas microbiológicas, descobriu, em 1882, o agente causador da tuberculose, nomeado, por isso, como de Bacilo de Koch.

Importante!!! Existem diferenças entre os termos endemia, epidemia e pandemia.

Endemia: é a ocorrência habitual de uma doença ou de um agente infeccioso, em determinada área geográfica; pode significar, também, a prevalência usual de determinada doença nessa área.

Epidemia: é doença geralmente infecciosa, de caráter transitório, que ataca simultaneamente grande número de indivíduos em uma determinada localidade.

Pandemia: é epidemia de grandes proporções, atingindo grande número de pessoas em uma grande área geográfica (um ou mais continentes).(MARTINS, 2018)

Até a primeira metade do século XX, os diversos ramos da medicina eram condicionados ao conhecimento derivado da microbiologia, porém, já na segunda metade do mesmo século, ocorrem mudanças na concepção das doenças prevalentes. Novas descobertas mostraram que, além de doenças infecciosas, também poderiam existir doenças crônicas e degenerativas, causadoras de morbidade e mortalidade, deixando claro que os agentes microbianos não eram capazes de explicar a etiologia de todas as doenças, assim como o prognóstico delas. Diante disso, princípios ambientais, sociais, culturais, psicológicos, genéticos, dentre outros, foram incorporados na elaboração dos muitos conceitos básicos para epidemiologia. Dessa forma, a epidemiologia começa, a partir da segunda metade do século XX, a se preocupar com as doenças crônicas não transmissíveis, tais como problemas cardíacos e câncer, sobretudo nos países industrializados.

A epidemiologia atual é uma disciplina relativamente nova e usa métodos quantitativos para estudar a ocorrência de doenças nas populações humanas e para definir estratégias de prevenção e controle.

Atualmente, as principais aplicações da epidemiologia são:

- ✓ Diagnóstico da situação de saúde em determinada localidade;
- ✓ Planejamento e organização dos serviços de saúde;
- ✓ Avaliação das tecnologias, programas ou serviços prestados na área da saúde;
- ✓ Aprimoramento na descrição do quadro clínico das doenças;
- ✓ Identificação de síndromes e classificação de doenças;
- ✓ Investigação da causa das doenças;
- ✓ Determinação de riscos;
- ✓ Determinação das probabilidades de adoecer;
- ✓ Determinação de prognósticos;
- ✓ Análise crítica de trabalhos científicos.

PROCESSO SAÚDE-DOENÇA

A distribuição das doenças na população é influenciada pelos aspectos biológicos dos indivíduos, pelos aspectos socioculturais e econômicos de sua comunidade e pelos aspectos ambientais do seu entorno, fazendo com que o processo saúde-doença se manifeste de forma diferenciada entre as populações (GOMES, 2015).

Processo saúde-doença é o termo utilizado para definir todas as variáveis envolvidas no estado de “saúde” e “doença” de um indivíduo ou população, levando-se em conta que ambos os estados estão interligados e que são consequências dos mesmos fatores. O termo “processo saúde-doença” se refere ao modo pelo qual ocorre, nos grupos da coletividade, o processo biológico de desgaste e reprodução, destacando como momentos particulares a presença de um funcionamento biológico diferente, com consequências para o desenvolvimento regular das atividades cotidianas, isto é, o surgimento de doença (GOMES, 2015).

HISTÓRIA NATURAL DA DOENÇA

A partir da definição de processo saúde-doença, devemos ter em mente que ele constitui uma das etapas pelas quais passa o indivíduo, ou a população, durante o processo de adoecimento, levando-se em consideração todas as variáveis que influenciam a saúde e as doenças, bem como seus desfechos, a cura ou a morte (GOMES, 2015).

Nessa perspectiva, a concepção de história natural da doença se torna fundamental. Um dos conceitos clássicos define história natural da doença como um conjunto de processos interativos que compreendem as inter-relações do agente etiológico, do susceptível e do meio ambiente, passando desde as variações ambientais/biológicas, que criam o estímulo patógeno, até a resposta do susceptível a esse agente, e que pode levar o indivíduo à doença, à invalidez, à recuperação ou à morte (GOMES, 2015).

A história natural da doença pode ser contada em dois períodos sequenciais: a vertente epidemiológica que compreende as relações susceptível e ambiente e a vertente patológica composta pelas alterações que ocorrem no organismo, após contato com os agentes causais. É um instrumento importante para o embasamento da compreensão de situações de adoecimento e as consequentes ações preventivas (GOMES, 2015).

A vertente epidemiológica corresponde ao período pré-patogênico. Essa é a fase em que ainda não há resposta biológica do organismo, a qual precisa ser bem conhecida para poder propor medidas de prevenção; daí a importância de se procurar a etiologia, identificando as relações causais, bem como as características dos agentes, dos fatores de risco, da intensidade da exposição, da susceptibilidade do organismo diante das agressões do local de ocorrência das situações de adoecimento (GOMES, 2015).

A vertente patológica corresponde ao período da patogênese. Essa é a fase em que já há resposta biológica do organismo, mas o conhecimento da etiologia é importante para a

definição de critérios diagnósticos e tratamento, visando à detecção e interrupção da evolução e regressão da doença instalada. No meio externo, interatuam os determinantes e os agentes causais das doenças e, no meio interno, ocorrem as modificações bioquímicas, fisiológicas e histológicas (GOMES, 2015).

As medidas preventivas dizem respeito a ações que evitem as doenças e os resultados delas. Podem ser classificadas em período pré-patogênico e período patogênico, sendo:

Período pré-patogênico: momento em que pode ocorrer a evolução das relações dinâmicas entre os condicionantes ecológicos; socioeconômicos e culturais; condições intrínsecas do sujeito.

Período patogênico: momento em que há interação agente-sujeito, provocando alterações bioquímicas, histológicas e fisiológicas, que vão evoluir para os sinais e os sintomas e, na sequência, para a cura, morte ou cronicidade.

Os níveis de prevenção podem ser divididos em:

Prevenção primária: promoção à saúde: diminuir risco de doenças - estabelecimento de padrões de vida, sociais, econômicos e culturais. Exemplos: moradia e alimentação adequadas; educação; áreas de lazer.

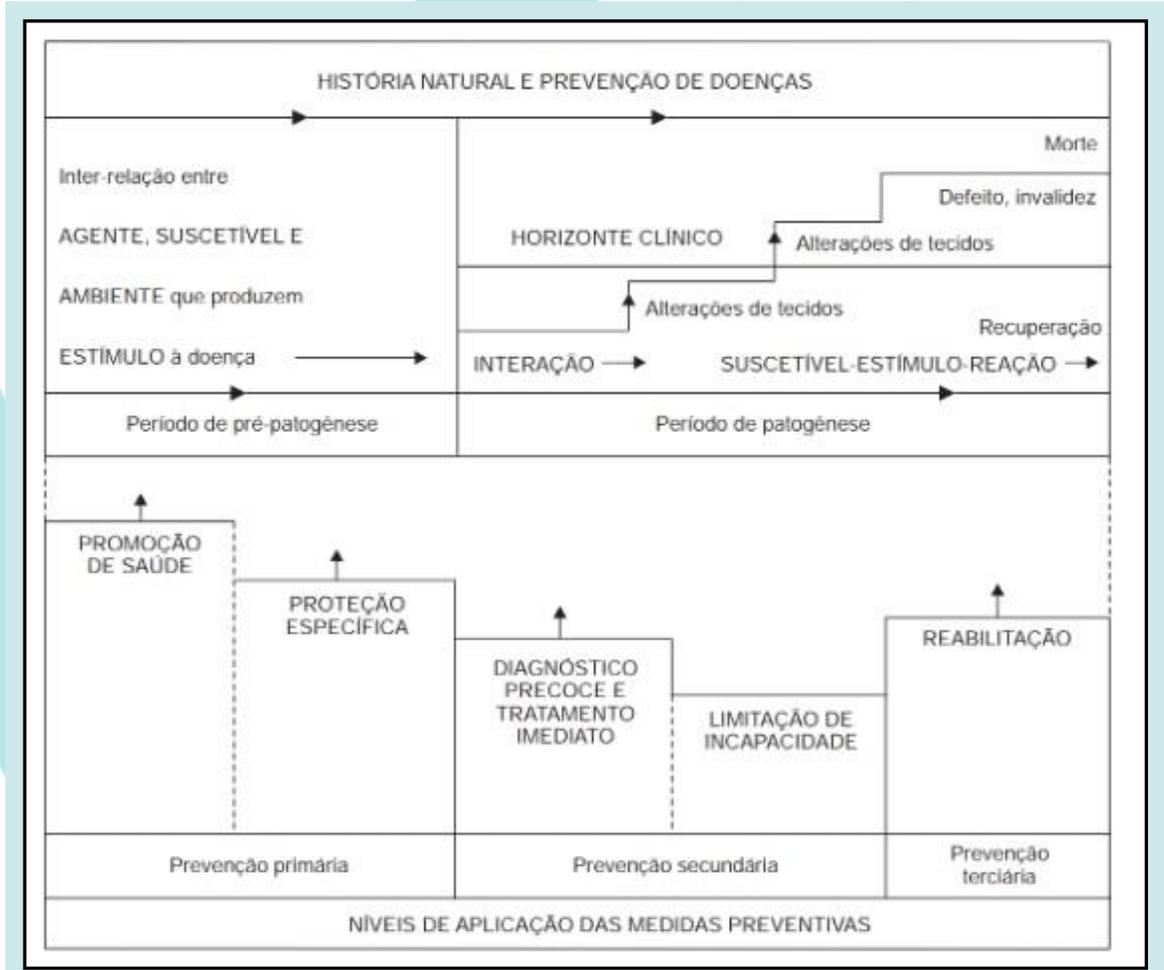
Proteção Específica: limitar incidência de doenças - controle de causas e fatores de risco. Exemplos: imunização, saúde ocupacional, higiene, proteção contra acidentes, aconselhamento genético e controle de vetores.

Prevenção secundária: visa curar e reduzir consequências da doença. É dirigida para o período entre o início da doença e o momento em que normalmente seria feito o diagnóstico.

Diagnóstico precoce: inquéritos para descoberta de casos na comunidade; exames periódicos, individuais, para detecção precoce dos casos. Tratamento imediato/limitação da incapacidade: evitar progressão da doença e sequelas.

Prevenção terciária: medidas para diminuir sofrimento provocado pela doença e promover adaptações às doenças incuráveis, reabilitação. Exemplos: fisioterapia, terapia ocupacional, emprego para o reabilitado. A FIG. 5 mostra o esquema de história natural de doença e níveis de prevenção.

FIGURA 5 - HISTÓRIA NATURAL E PREVENÇÃO DE DOENÇAS



FONTE 5:ROUQUAYROL, 2017, P. 19.

AS FASES DO MÉTODO ESTATÍSTICO

Em um estudo estatístico, normalmente, segue-se um conjunto de passos que designamos por fases do método estatístico, como podemos ver na FIG. 6.

FIGURA 6 - ESQUEMA DAS FASES DO MÉTODO ESTATÍSTICO



FONTE 6: ELABORAÇÃO PRÓPRIA, 2023.

Definição do problema: A primeira fase consiste na definição e na formulação correta do problema a ser estudado. Pode responder à definição de um problema ou, simplesmente, dar resposta a um interesse dos alunos. Em alguns casos, estão envolvidas variáveis qualitativas e, em outros, variáveis quantitativas.

Planejamento: Uma vez que o problema já foi definido, é preciso determinar um processo para o resolver, em especial, a forma como obter informações sobre a variável ou variáveis em estudo. É nessa fase que se decide pela observação da população ou de uma amostra.

Coleta de dados: Nessa fase, define-se se os dados serão de origem primária (coletados de forma direta) ou secundários (coletados indiretamente), além de decidir os

instrumentos utilizados para obtenção dos dados como, por exemplo, questionários, observação, experimentação, pesquisa bibliográfica, dentre outros.

Organização de dados: Consiste em “resumir” os dados através da sua contagem e agrupamento em tabelas estatísticas. Assim, obtém-se um conjunto de informações que irá conduzir ao estudo.

Apresentação de dados: Em estatística, há duas formas de apresentação dos dados que não se excluem mutuamente e que são fundamentais para a fase seguinte: apresentação por tabelas e/ou apresentação por gráficos. Essas formas de apresentar permitem sintetizar grandes quantidades de dados, tornando mais fácil a compreensão da variável em estudo e permitindo a análise a ser realizada na próxima fase do método.

Análise e interpretação de dados: Nessa fase, calculam-se novos números com base nos dados estatísticos. Esses novos números permitem fazer uma descrição do fenômeno, evidenciando algumas das suas características particulares. Através da análise quantitativa e qualitativa dos dados é possível estabelecer relações, validar hipóteses e fazer inferências.

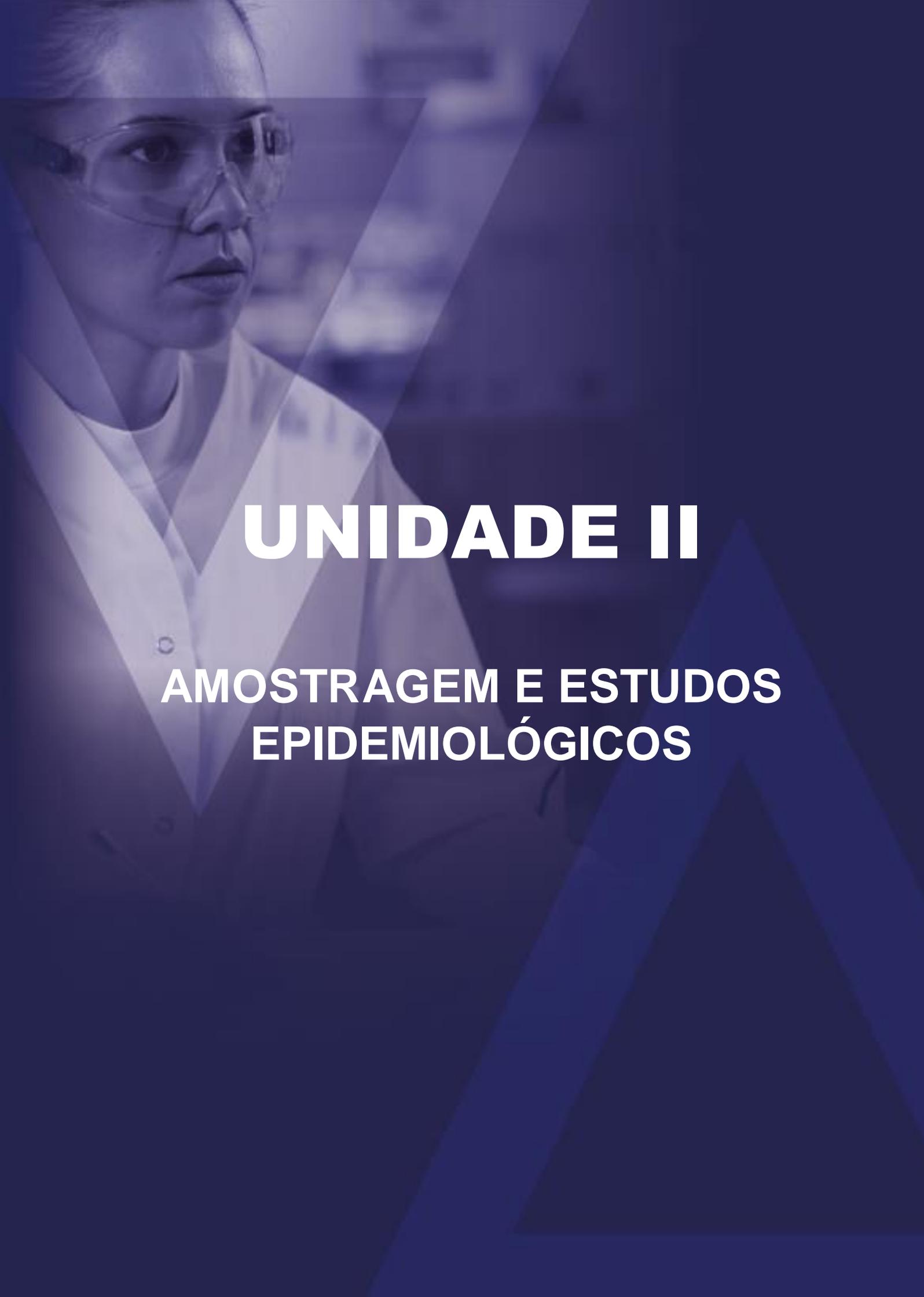


RESUMO DA UNIDADE

Nesta unidade aprendemos sobre os conceitos iniciais de estatística, nos quais população e amostra são fundamentais para definição dos conjuntos de dados. Os dados, quando obtidos diretamente, constituem o que chamamos de dados primários e são originados de fontes primárias e, quando obtidos por meios de banco de dados e outras fontes de pesquisas já realizadas são fontes secundárias e os dados são chamados de dados secundários. As variáveis são classificadas quanto a sua natureza e subdivididas em tipos quanto a sua apresentação; dessa forma, temos as variáveis de natureza quantitativa, sendo elas contínuas ou discretas e de natureza qualitativa, sendo do tipo nominal ou ordinal.

A epidemiologia é uma ciência que estuda o processo saúde-doença em coletividades humanas, analisando a distribuição e os fatores determinantes das enfermidades, danos à saúde e eventos associados à saúde coletiva, propondo medidas específicas de prevenção, controle ou erradicação de doenças e fornecendo indicadores que sirvam de suporte ao planejamento, administração e avaliação das ações de saúde.

A epidemiologia reúne métodos e técnicas de três áreas principais de conhecimento: estatística, ciências biológicas e ciências sociais. A área de atuação da epidemiologia é bastante ampla. Em síntese, pode-se afirmar que a distribuição das doenças nas populações é influenciada pelos aspectos biológicos dos indivíduos, pelos aspectos socioculturais e econômicos de sua comunidade e pelos aspectos ambientais do seu entorno, fazendo com que o processo saúde-doença se manifeste de forma diferenciada entre elas.



UNIDADE II

AMOSTRAGEM E ESTUDOS EPIDEMIOLÓGICOS



OBJETIVOS:

- ✓ Analisar os tipos de amostras probabilísticas e suas aplicações.
- ✓ Discutir a representatividade de amostras e suas formas de cálculo amostral.
- ✓ Descrever os tipos de estudos epidemiológicos e suas aplicações.



META DA UNIDADE

O maior desafio que o profissional de saúde enfrenta na realização de um estudo de natureza epidemiológica é reunir, apropriadamente, uma quantidade suficiente de indivíduos para compor o(s) grupo(s) a ser(em) investigado(s). Alguns questionamentos que surgem nesse momento são: todos os indivíduos da população devem ser incluídos na pesquisa ou apenas uma amostra? No caso de a escolha recair sobre a amostra, como se pode obtê-la de maneira correta? Quantas pessoas são necessárias para formá-la? (PEREIRA, 1995).

A teoria sobre amostragem estatística fornece técnicas sofisticadas e confiáveis para que a seleção das amostras seja feita da forma adequada e representativa, considerando-se a necessidade de lidar com os múltiplos fatores envolvidos. As fórmulas e as explicações técnicas, em nível matemático, encontradas nas publicações especializadas, assustam o profissional de saúde, em geral não familiarizado com esse nível de detalhamento. No entanto, para a maioria das situações enfrentadas pelos profissionais da área de saúde, apenas um conhecimento elementar do tema é suficiente, uma vez que o pesquisador já esteja totalmente convencido da necessidade de utilizar amostras de maneira correta, quando se faz necessário.

Estudaremos, aqui, as técnicas de amostragem probabilísticas mais empregadas em estudo na área da saúde humana e animal, bem como elementos fundamentais para o sucesso da utilização de amostras em uma pesquisa.

Vamos, ainda, estudar os principais delineamentos de pesquisa epidemiológica, os delineamentos epidemiológicos que compreendem ensaios experimentais e não experimentais, os quais, por sua vez, subdividem-se em tipos que veremos com mais detalhes e exemplos.

TIPOS DE AMOSTRAS PROBABILÍSTICAS

As amostras aleatórias, também chamadas de “casuais”, “probabilísticas”, “estatísticas” ou “ao acaso”, são empregadas quando a finalidade é determinar as características da população a partir do exame de um pequeno número de unidades. Elas são solicitadas para situações em que as variações individuais têm de ser levadas em conta. As unidades para essa amostra são escolhidas ao acaso, de modo a obter uma amostra que tenha as mesmas características da população.

Conceito de seleção ao acaso (ou aleatória) - Vamos conceituar a expressão “ao acaso” do ponto de vista estatístico, que nesse contexto, é a noção central do processo de amostragem. Aleatório ou ao acaso quer dizer que cada unidade tem a mesma probabilidade de fazer parte da amostra, ou seja, todos os elementos têm a mesma chance de serem escolhidos. Veja o seguinte exemplo.

EXEMPLO

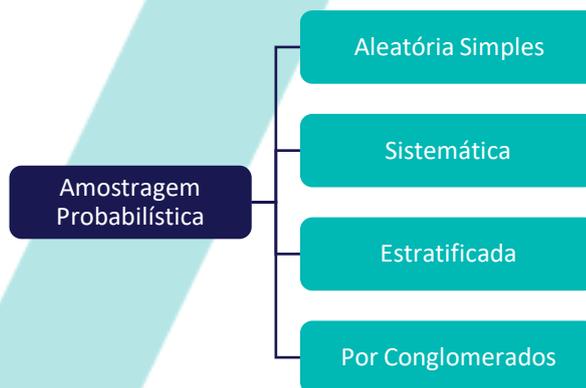
Amostragem por telefone:

Como se sabe, no país, nem todas as residências são providas de telefone. Logo, compor uma amostra a partir dos nomes que constam da lista telefônica não alcançará o objetivo de constituir amostra ao acaso, representativa de toda a população. Os que não possuem telefone não têm chance de aparecer na amostra. Por isso, não se faz pesquisa por telefone para coletar informações quantitativas sobre a população geral, mas a técnica pode ser útil para segmentos da população, como os membros de um clube recreativo de classe média ou de uma associação profissional ou, como já referido, para obter informações gerais sobre um dado evento (PEREIRA, 1995).

O profissional de saúde deve, portanto, estar atento aos significados da palavra “acaso”, que tem diferente conotação para o leigo e para o especialista. Para obter amostras ao acaso, de natureza estatística, todo julgamento humano é excluído. Somente nessas condições podem-se usar as regras de inferência estatística (cálculo do erro amostral e determinação de frequências por intervalo), de modo a fazer generalizações para a população a partir dos dados da amostra (VIEIRA, 2018).

Há diversas maneiras de selecionar indivíduos para formar uma amostra probabilística, dependendo das características da população que se deseja conhecer (FIG. 1). Os princípios básicos aqui apresentados se aplicam igualmente a outras populações, como de animais e plantas.

FIGURA 7 - TIPOS DE AMOSTRAGEM PROBABILÍSTICA



FONTE 7: ELABORAÇÃO PRÓPRIA, 2023

O modo de seleção das pessoas para compor a amostra, quando se busca torná-la representativa da população, varia desde a forma aleatória simples e a sistemática, de escolha, a esquemas mais elaborados, nos quais são usados a estratificação e os conglomerados, em vários estágios.

AMOSTRAGEM ALEATÓRIA SIMPLES

Esse método é uma forma elementar de procedimento. Para sua preparação, há necessidade de uma lista das unidades a serem sorteadas. Essa lista de unidade é o “marco” ou “cadastro” de amostragem.

O procedimento pode ser o seguinte: numeram-se todos os indivíduos que compõem a população e realiza-se um sorteio das unidades, por qualquer método, inclusive computador (loteria), ou usando quadro de números aleatórios (que é uma loteria já realizada).

Você, com certeza, já participou e/ou já realizou algo semelhante a esse método no seu cotidiano, para saber, por exemplo, quem será encarregado de determinadas tarefas: o nome de cada um é colocado em um pedaço de papel, dobrado e juntado aos demais. Vários nomes podem ser, dessa maneira, escolhidos, ao acaso, para compor a amostra. Essa técnica funciona bem para populações pequenas. Uma outra forma de simplificar o processo de seleção é utilizando uma tabela de números aleatórios (VIERIA, 2018).



INDICAÇÃO DE VÍDEO

Veja neste vídeo como funciona a **Amostragem Aleatória Simples – AAS**.

https://youtu.be/Tk-7ftUP_Y

AMOSTRAGEM SISTEMÁTICA

É uma forma também elementar de procedimento e vista como uma variação da AAS, conveniente quando a população está ordenada segundo algum critério, como fichas de um fichário, listas telefônicas etc. Sua realização se dá em algumas etapas bem simples:

Passo 1: calcula-se o intervalo (k) de amostragem $k = \frac{N}{n}$, aproximando-o para o inteiro mais próximo, em que N é o tamanho da população e n , o tamanho da amostra.

Passo 2: utilizando-se uma tabela de números aleatórios, sorteia-se um número x entre 1 e k , formando, assim, a amostra dos elementos correspondente aos números: x ; $x + k$; $x + 2k$;
Por exemplo, para selecionar 200 radiografias de uma série de 1000, podemos proceder da seguinte forma: $N = 1.000$, $n = 200$. Logo: $k = \frac{1000}{200} = 5$

Supondo que 3 seja o número sorteado entre 1 e 5., os elementos (radiografias) da população numerados por 3, 8, 13, ..., 998 irão compor a amostra.

Embora esse tipo de amostragem seja fácil de ser conduzido e economicamente acessível, sua desvantagem é dada pela dificuldade de obter as distribuições amostrais das estatísticas com os dados amostrais, comprometendo a qualidade das inferências sobre a população. Podemos exemplificar isso a partir da situação relatada no exemplo anterior. Para tanto, vamos supor que, no arquivo radiográfico de um certo hospital, o funcionário responsável tenha resolvido classificar as radiografias segundo sua procedência. Assim, atribuiu terminação par a todos os exames provenientes da enfermaria de pneumologia e “distribuiu” os demais algarismos pelas demais enfermarias. Nessa circunstância, caso o pesquisador não tenha conhecimento do “método”, certamente incorreria em vício amostral ao sistematizar sua seleção segundo o exemplo anterior, causando distorção no processo sistemático de amostragem (VIEIRA, 2018).



INDICAÇÃO DE VÍDEO

Veja neste vídeo como funciona a **Amostragem Sistemática**.

<https://youtu.be/H1jq8bEvl34>

AMOSTRAGEM ESTRATIFICADA

No caso de população heterogênea, em que se podem distinguir subpopulações mais ou menos homogêneas, denominadas estratos, é possível utilizar o processo de amostragem estratificada. Às vezes, é conveniente a divisão da população em alguns estratos (grupos), seja de sexo, idade, nível socioeconômico ou outro atributo, caracterizando a amostra estratificada. Optando por esse procedimento, separa-se a população em subgrupos que não se superponham e escolhe-se uma amostra aleatória simples ou sistemática de cada estrato. É técnica recomendada para garantir a presença de determinados segmentos da população que, de outra maneira, poderiam estar sub-representações ou ausentes (TRIOLA, 2017).

Imagine que você quer obter uma amostra dos funcionários de um hospital. Essas pessoas são diferentes em relação ao sexo, à faixa de idade, à função. Constituem, portanto, uma população heterogênea. Você não deve selecionar uma amostra aleatória dessa população, porque corre o risco de obter uma amostra pouco representativa como, por exemplo, uma amostra constituída apenas por enfermeiros. Então, o que deve ser feito?

Antes de coletar a amostra, reúna os nomes das pessoas de mesmo sexo, mesma faixa de idade e mesma função em grupos homogêneos denominados, tecnicamente, estratos. Depois, colete uma amostra aleatória dentro de cada estrato. Junte essas amostras para formar uma só, que será a amostra estratificada.

Existem, porém, dois tipos de amostra estratificada: a proporcional e a não proporcional. Na amostra estratificada proporcional, o tamanho de cada estrato é proporcional ao tamanho do estrato na população. Em outras palavras, a proporção de unidades de cada estrato na amostra é igual à proporção de cada estrato na população. Veja os exemplos a seguir:

EXEMPLO

Amostra estratificada proporcional: imagine que você tem 500 cadastros arquivados em sua empresa, 300 de homens e 200 de mulheres. Você quer obter dados dos dois sexos. O que você faria?

Solução: Comece separando os cadastros de homens dos cadastros de mulheres. Você tem, então, dois estratos, um de homens, outro de mulheres. Para obter uma amostra estratificada proporcional de dez cadastros, ou seja 20%, tire uma amostra casual de seis cadastros de homens (que corresponde a 20% do total de homens) e quatro cadastros de mulheres (que corresponde a 20% do total de mulheres). Reúna os dados dos dois estratos em uma só amostra estratificada.

EXEMPLO

Amostra estratificada não proporcional: imagine que você tem 500 cadastros arquivados em sua empresa, 300 de homens e 200 de mulheres. Você quer obter dados dos dois sexos. O que você faria?

Solução: Comece separando os cadastros de homens dos cadastros de mulheres. Você tem, então, dois estratos, um de homens, outro de mulheres. Se você optar por uma amostra estratificada não proporcional de dez cadastros, tire uma amostra casual de cinco cadastros de homens e cinco de mulheres. Reúna os dados dos dois estratos em uma só amostra estratificada.



INDICAÇÃO DE VÍDEO

Veja neste vídeo como funciona a Amostragem Estratificada.

<https://youtu.be/PNdMAnE96cU>

AMOSTRAGEM POR CONGLOMERADOS

Em algumas situações, o fato de a população ser muito dispersa, no espaço ou no tempo, acarreta enorme dificuldade para o uso das amostragens aleatória simples, sistemática e a estratificada. É o caso de seleção de amostra de um país, estado ou cidade. Há formas de minimizar a extensão do trabalho de campo pela escolha de alguns grupos ou áreas apenas, denominados aqui de “conglomerados”. O procedimento é o seguinte:

Passo 1: faz-se a divisão da população em conglomerados que não se superponham.

Passo 2: selecionam-se alguns conglomerados, ao acaso; somente os conglomerados sorteados compõem a amostra. Não há realmente interesse em comparar esses conjuntos entre si, mas usá-los, somados, como amostra representativa de todo o universo.

Veja o seguinte exemplo: peso ao nascer em maternidades.

Para compor uma amostra de recém-nascidos de uma grande cidade, onde ocorrem, anualmente, dezenas de milhares de partos, poderíamos, inicialmente, fazer uma lista de todas as maternidades; em seguida, selecionar algumas maternidades ao acaso e nelas, por levantamento dos dados nos prontuários, recolher as informações sobre o peso ao nascer, no ano considerado (TRIOLA, 2017).



INDICAÇÃO DE VÍDEO

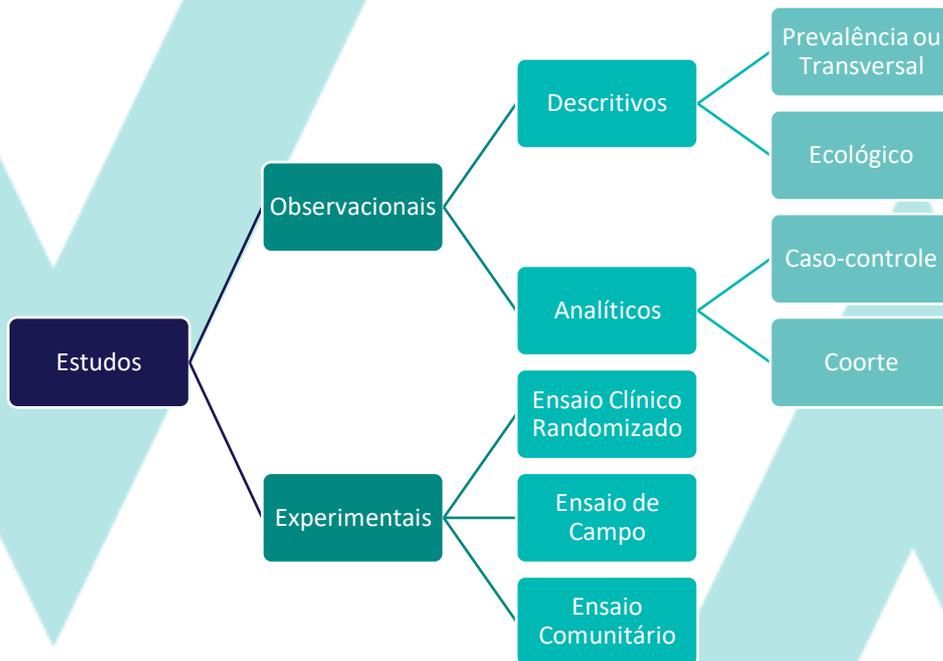
Veja neste vídeo como funciona a Amostragem Conglomerados.

<https://youtu.be/Elf8x6vucRs>

PRINCIPAIS TIPOS DE ESTUDOS EPIDEMIOLÓGICOS EXPERIMENTAIS E OBSERVACIONAIS

Os estudos epidemiológicos podem ser classificados em observacionais ou experimentais. Os tipos mais comuns de estudos estão listados na FIG. 2.

FIGURA 8 - ESQUEMA DOS PRINCIPAIS TIPOS DE ESTUDOS EPIDEMIOLÓGICOS



FONTE 8: ELABORAÇÃO PRÓPRIA, 2023

ESTUDOS OBSERVACIONAIS

Os estudos observacionais permitem que a natureza determine o seu curso: o investigador mede, mas não intervém. Esses estudos podem ser descritivos e analíticos:

- Um **estudo descritivo** se limita a descrever a ocorrência de uma doença em uma população, sendo, frequentemente, o primeiro passo de uma investigação epidemiológica;
- Um **estudo analítico** aborda, com mais profundidade, as relações entre o estado de saúde e as outras variáveis.

ESTUDOS DESCRITIVOS

Estudo de Prevalência ou Transversal

Esse tipo de estudo consiste em estudar determinada população, ou uma amostra representativa dela, em função de apresentar características que possibilitem a investigação exposição-doença. Nesse tipo de estudo, as informações sobre exposições e doenças da população são medidas, simultaneamente, em um curto período de tempo. Assim, o estudo transversal fornece um retrato de como as variáveis estavam relacionadas no momento da pesquisa.

Em outras palavras, os estudos transversais medem a prevalência da doença e, por essa razão, são frequentemente chamados de estudos de prevalência. Em um estudo transversal, as medidas de exposição e efeito (doença) são realizadas ao mesmo tempo. Por esse motivo, não é fácil avaliar as associações encontradas nesses estudos. A questão-chave nesse tipo de delineamento é saber se a exposição precede ou é consequência do efeito. Se os dados coletados representam a exposição antes da ocorrência de qualquer efeito, a análise pode ser feita de modo semelhante à utilizada nos estudos de coorte (MARTINS, 2018).

Os estudos transversais são relativamente baratos, fáceis de conduzir e úteis na investigação das exposições que são características individuais fixas, tais como grupo étnico e grupo sanguíneo. Na investigação de surtos epidêmicos, a realização de um estudo transversal medindo diversas exposições é, em geral, o primeiro passo para a determinação da sua causa (ROUQUAYROL, 2017).

ESTUDO ECOLÓGICO

Nesse delineamento, a unidade de análise não é constituída de indivíduos, mas de grupos de indivíduos. Assim, não se sabe se um determinado indivíduo da população investigada é exposto ou doente, apenas as informações globais estão disponíveis, como a proporção ou frequência da doença na população e a proporção ou frequência da exposição na população. Os estudos ecológicos (ou de correlação) são úteis para gerar hipóteses.

EXEMPLO

Por exemplo, foi encontrada uma associação entre média de vendas de droga antiasmática e a ocorrência de um número elevado de óbitos por asma em diferentes províncias da Nova Zelândia. Tais observações deveriam ser testadas levando-se em conta todos os potenciais fatores de confusão a fim de excluir a possibilidade de que outras características, por exemplo, a severidade da doença em diferentes populações, não seja a responsável por essa associação. (ROUQUAYROL,2017)

ESTUDOS ANALÍTICOS

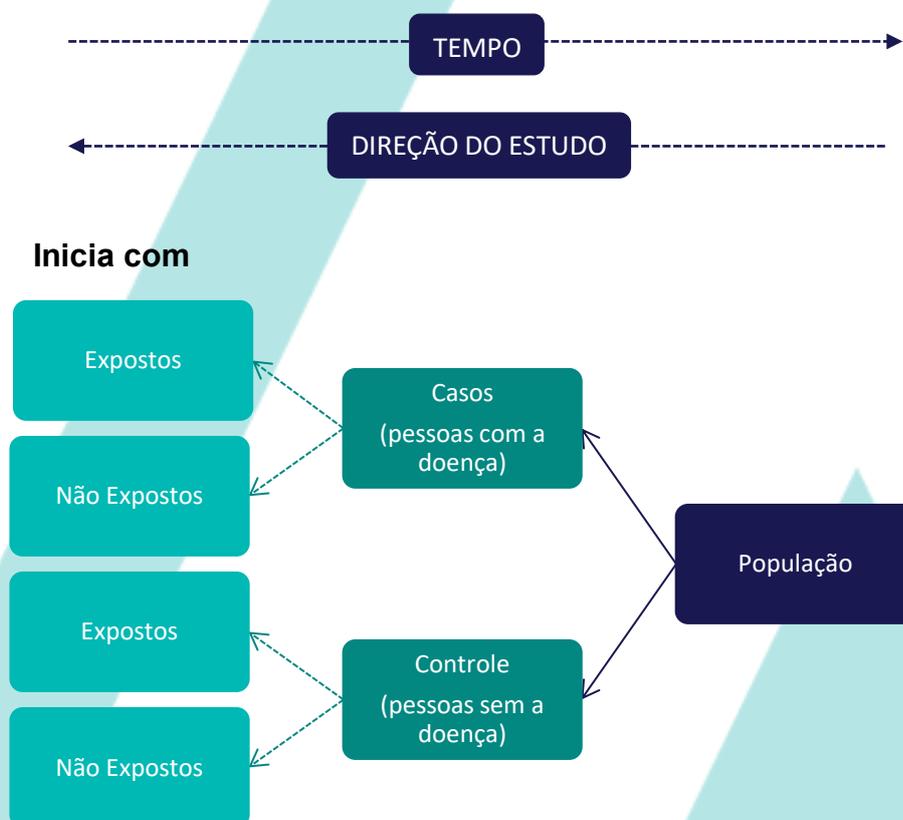
Estudo retrospectivo ou de casos e controles

Estudos de casos e controles constituem uma forma relativamente simples de investigar a causa das doenças, particularmente doenças raras. Esse tipo de estudo inclui pessoas com a doença (ou outra variável de desfecho) e um grupo controle (grupo de comparação ou referência) composto de pessoas não afetadas pela doença ou variável de desfecho. A ocorrência de uma possível causa é comparada entre casos e controles. Os investigadores coletam dados sobre a ocorrência da doença em um determinado momento no tempo e sobre a ocorrência de exposições em algum momento no passado (MARTINS, 2018).

Os estudos de casos e controles são, portanto, longitudinais, diferentes dos estudos transversais (FIG. 3). Os estudos de casos e controles também são chamados de retrospectivos, uma vez que o investigador busca, no passado, uma determinada causa (exposição) para a doença ocorrida. Entretanto, isso pode causar confusão, porque os termos retrospectivos e prospectivos também são utilizados para descrever o tempo da coleta dos dados em relação ao momento atual. Nesse caso, um estudo de casos e controles pode ser tanto retrospectivo, quando os dados fazem referência ao passado, quanto prospectivo, quando os dados são continuamente coletados no decorrer do tempo (MARTINS, 2018).

Esquema de um estudo de caso-controle

FIGURA 9 - ESQUEMA DE UM ESTUDO DE CASO-CONTROLE



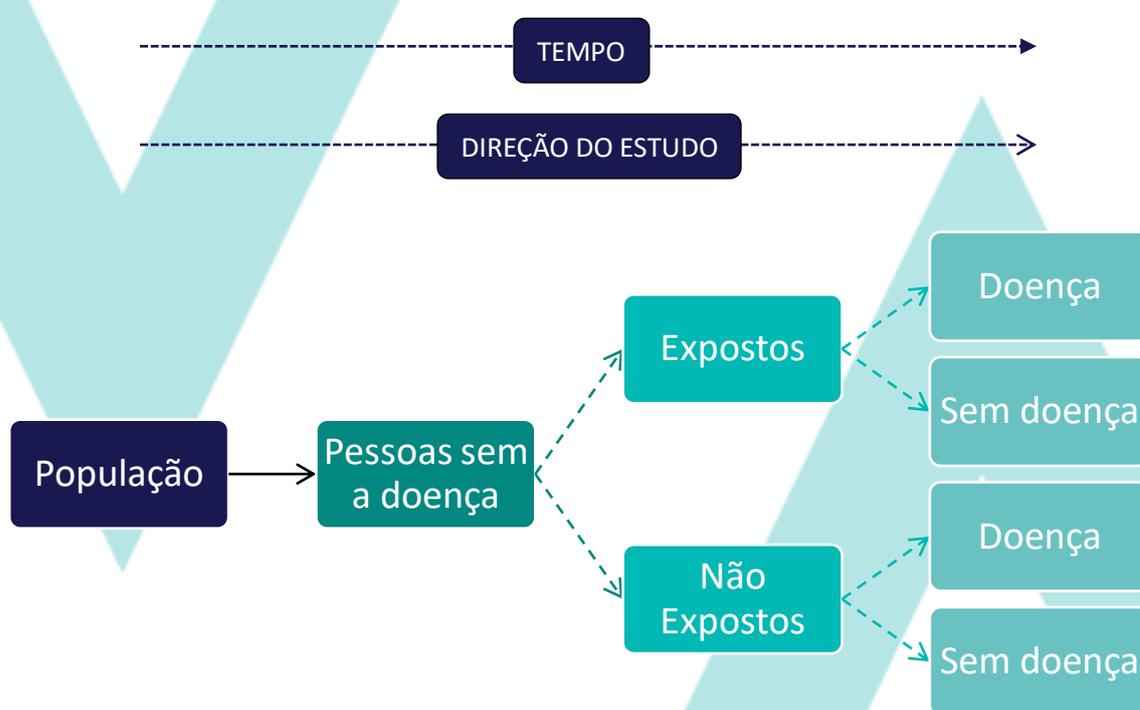
FONTE 9: ELABORAÇÃO PRÓPRIA, 2023

ESTUDO DE COORTES (GRUPOS)

Os estudos de coorte, também chamados longitudinais ou de incidência, iniciam com um grupo de pessoas livres da doença, classificadas em subgrupos, de acordo com a exposição a uma causa potencial da doença ou desfecho sob investigação (FIG. 4). As variáveis de interesse são especificadas e medidas e a coorte inteira acompanhada, com o objetivo de ver se o surgimento de novos casos de doença (ou outro desfecho) difere entre os grupos, conforme a presença ou não de exposição. Em virtude de os dados coletados fazerem referência a diferentes pontos no tempo, os estudos de coorte são longitudinais, fato que os torna semelhantes aos de casos e controles. Os estudos de coorte são chamados de estudos prospectivos, mas essa terminologia é confusa e, por isso, deve ser evitada (ROUQUAYROL, 2017).

Como mencionado previamente, o termo “prospectivo” se refere ao momento da coleta dos dados e não à relação entre exposição e efeito (doença). Assim, os estudos de coorte podem ser tanto prospectivos quanto retrospectivos. Os estudos de coorte fornecem a melhor informação sobre a etiologia das doenças e a medida mais direta do risco de desenvolvê-la. Embora conceitualmente simples, os estudos de coorte são bastante caros, porque podem requerer longos períodos de acompanhamento, visto que a doença pode ocorrer após uma exposição prolongada. Por exemplo, o período de indução para a leucemia causada por radiação (isto é, o tempo necessário para uma causa específica produzir um desfecho) é de

FIGURA 10 - ESQUEMA DE UM ESTUDO DE COORTE



FONTE 10: ELABORAÇÃO PRÓPRIA, 2023

muitos anos, sendo, portanto, necessário acompanhar os participantes por um longo período (ROUQUAYROL, 2017).

ESTUDOS EXPERIMENTAIS

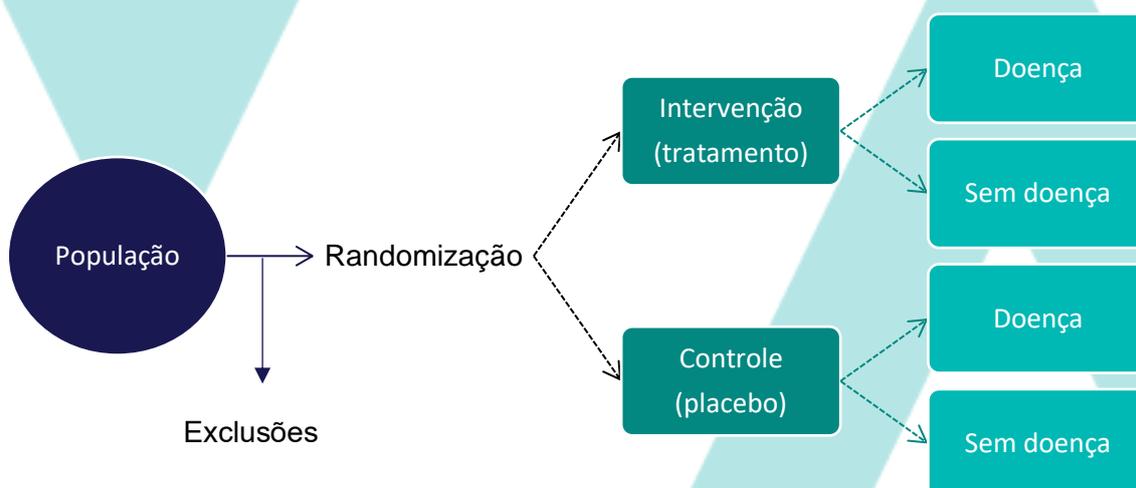
Estudos experimentais ou de intervenção envolvem a tentativa de mudar os determinantes de uma doença, tais como uma exposição ou comportamento, ou cessar o progresso de uma doença através de tratamento. São similares a experimentos realizados em outras ciências.

ENSAIO CLÍNICO ALEATÓRIO

O ensaio clínico randomizado é um experimento epidemiológico que tem por objetivo estudar os efeitos de uma intervenção em particular. Os indivíduos selecionados são aleatoriamente alocados para os grupos intervenção e controle, e os resultados são avaliados comparando-se os desfechos entre esses grupos.

O diagrama esquemático de um ensaio clínico randomizado é mostrado na FIG. 5. Para assegurar que os grupos comparados sejam equivalentes, os pacientes são alocados aleatoriamente, ou seja, ao acaso. Isso garante a comparabilidade entre os grupos intervenção e controle desde o início da intervenção. Assim, quaisquer diferenças observadas entre eles serão decorrentes do acaso, não sendo, portanto, afetadas por viés do investigador (ROUQUAYROL, 2017).

FIGURA 11 - ESQUEMA DE ENSAIO CLÍNICO RANDOMIZADO



FONTE 11: ELABORAÇÃO PRÓPRIA, 2023

ENSAIOS DE CAMPO

Ensaio de campo, em contraste com os ensaios clínicos, envolvem pessoas que estão livres de doença, mas sob risco de desenvolvê-la. Os dados são coletados “no campo”, usualmente entre pessoas da população geral não institucionalizadas. Uma vez que os participantes estão livres da doença e o propósito é prevenir a ocorrência de doenças, mesmo entre aquelas de baixa frequência, os ensaios de campo envolvem um grande número de pessoas, o que os torna caro e logisticamente complicados. Os ensaios de campo podem ser

utilizados para avaliar intervenções que objetivam reduzir a exposição, sem necessariamente medir a ocorrência dos efeitos sobre a saúde.

ENSAIOS COMUNITÁRIOS

Nesse tipo de experimento, os grupos de tratamento são comunidades ao invés de indivíduos. Esse delineamento é particularmente apropriado para doenças que tenham suas origens nas condições sociais e que possam ser facilmente influenciadas por intervenções dirigidas ao comportamento do grupo ou do indivíduo. Uma limitação desse tipo de delineamento é que somente um pequeno número de comunidades pode ser incluído e a alocação aleatória das comunidades não é muito prática. Assim, outros métodos são requeridos para assegurar que quaisquer diferenças encontradas ao final do estudo possam ser atribuídas à intervenção e não a diferenças inerentes às comunidades. Além disso, é difícil isolar as comunidades nas quais a intervenção está sendo conduzida devido a mudanças sociais em curso (ROUQUAYROL, 2017).



RESUMO DA UNIDADE

Amostragem Probabilística é aquela em que a seleção é aleatória, de tal forma que cada elemento da população tem uma probabilidade conhecida de fazer parte da amostra. Como tipo de amostragem probabilística vimos: Amostragem Aleatória Simples (AAS), indicada para populações homogêneas; Amostragem Sistemática, na qual a população deve ser ordenada de forma que os elementos sejam identificados pela posição; Amostragem Estratificada, que consiste em dividir a população em subgrupos mais homogêneos (estratos), de tal forma que haja uma homogeneidade dentro dos estratos e uma heterogeneidade entre os estratos e Amostragem por Conglomerados, método muito utilizado por motivos de ordem prática e econômica, em que se divide uma população em pequenos grupos e se sorteia um número suficiente desses pequenos grupos (conglomerados), cujos elementos constituirão a amostra.

Os estudos epidemiológicos incluem vigilância, observação, pesquisa analítica e experimento. O objetivo geral do estudo da epidemiologia é reduzir os problemas de saúde da população. Na prática, ela estuda principalmente a ausência de saúde sob as formas de doenças e seu nível de gravidade e complexidade.

Os estudos descritivos têm por objetivo determinar a distribuição de doenças ou condições relacionadas à saúde, segundo o tempo, o lugar e/ou as características dos indivíduos, ou seja, precisam responder às perguntas: quando, onde e quem adoece? A epidemiologia descritiva pode fazer uso de dados primários (dados coletados para o desenvolvimento do estudo) e dados secundários (dados pré-existentes de mortalidade e hospitalizações, por exemplo).

Estudos analíticos são aqueles delineados para examinar a existência de associação entre uma exposição e uma doença ou condição relacionada à saúde.

Nos estudos ecológicos, compara-se a ocorrência da doença/condição relacionada à saúde e a exposição de interesse entre agregados de indivíduos (populações de países, regiões ou municípios, por exemplo), para verificar a possível existência de associação entre elas.

Os estudos caso-controle e os estudos de coorte podem ser utilizados para investigar a etiologia de doenças ou de condições relacionadas à saúde entre idosos, determinantes da longevidade, e para avaliar ações e serviços de saúde.

Nos estudos de coorte, primeiramente, identifica-se a população de estudo e os participantes são classificados em expostos e não expostos a um determinado fator de interesse. Nesse tipo de estudo, a mensuração da exposição antecede o desenvolvimento da doença, não sendo sujeita ao viés de memória como nos estudos caso-controle.



UNIDADE III

ESTATÍSTICAS DESCRITIVAS



OBJETIVOS

- ✓ Compreender as principais estatísticas descritivas básicas e formas de cálculo e aplicações.



METAS

As Unidades 1 e 2 estudadas até aqui apresentaram alguns conceitos importantes, incluindo o contexto e a fonte dos dados, o método de amostragem, as conclusões e as implicações práticas. É comum que muitas amostras de dados sejam grandes, de modo que, para compreendê-las, devemos organizá-las, resumi-las e representá-las de forma que permitam ser mais bem entendidas. Podemos organizar e resumir dados numericamente em tabelas, ou visualmente, em gráficos, e é o que iremos abordar nesta Unidade 3. Entretanto, nosso objetivo final não é apenas gerar uma tabela ou um gráfico, mas, sobretudo, desejamos utilizar as tabelas e os gráficos como ferramentas que expõem características escondidas e importantes dos dados. Nesta unidade, estamos focados, principalmente, na distribuição de um conjunto de dados, o qual consiste em uma das cinco características a seguir que, tipicamente, são as mais importantes.

Características dos Dados:

1. **Centro:** um valor representativo que indica a localização do meio do conjunto de dados.
2. **Varição:** uma medida de quanto variam os valores dos dados.
3. **Distribuição:** a natureza ou forma do espalhamento dos dados no conjunto de valores (como um sino).
4. **Outliers ou Valores Atípicos:** valores amostrais que se encontram distantes da maioria dos outros.
5. **Tempo:** qualquer mudança nas características dos dados ao longo do tempo.

Assim, aqui estamos de olho na distribuição dos dados, e posteriormente, discutiremos sobre os métodos para a investigação de outras características.

TABELAS DE FREQUENCIAS E GRÁFICOS

Ao se trabalhar com grandes conjuntos de dados, uma distribuição de frequência (ou tabela de frequência), em geral, é útil para organizá-los e resumi-los. Uma distribuição de frequência nos ajuda a entender a natureza da distribuição de um conjunto de dados. Sendo assim, podemos dar a seguinte definição.

DEFINIÇÃO

Uma distribuição de frequência (ou tabela de frequência) mostra como o conjunto de dados é dividido entre várias categorias (ou classes), listando as categorias juntamente com o número (frequências) de valores de dados em cada uma delas. (Triola, 2017)

EXEMPLO

Tabelas de distribuição de frequências podem ser utilizadas para a apresentação de dados categorizados, discretos e contínuos.

APRESENTAÇÃO DE DADOS CONTÍNUOS EM TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS

Considere os escores de QI (Quociente de Inteligência) do grupo de baixo nível de chumbo listados na Tabela 1. A Tabela 2 é uma distribuição de frequência que resume os escores de QI. A frequência de uma classe particular é o número dos valores originais que caem dentro daquela classe. Por exemplo, a primeira classe na Tabela 2 tem uma frequência de 2, o que indica que 2 dos escores de QI estão entre 50 e 69, inclusive.

TABELA 1 - ESCORES DE QI COMPLETOS DO GRUPO DE BAIXO NÍVEL DE CHUMBO E DO GRUPO DE ALTO NÍVEL DE CHUMBO

Baixo Nível de Chumbo (Grupo 1)															
70	85	86	76	84	96	94	56	115	97	77	128	99	80	118	86
141	88	96	96	107	86	80	107	101	91	125	96	99	99	115	106
105	96	50	99	85	88	120	98	87	98	78	100	87	87	94	89
80	111	104	85	94	75	73	76	107	88	89	96	97	97	76	107
10	85	76	95	86	89	76	96	101	108	102	77	92	92		
Alto Nível de Chumbo (Grupo 3)															
82	93	85	75	85	80	101	89	80	94	88	104	88	88	83	104
96	76	80	79	75											

FONTE: TRIOLA (2017, p. 42).

TABELA 2 - ESCORES DE QI COMPLETOS DO GRUPO DE BAIXO NÍVEL DE CHUMBO

Escore de QI	Frequência (<i>f_i</i>)
50 – 69	2
70 – 89	33
90 – 109	35
110 – 129	7
130 – 149	1

FONTE: ELABORAÇÃO PRÓPRIA (2023).

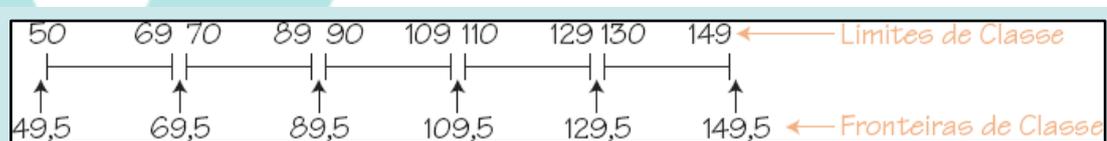
Definições importantes

Limites inferiores de classe são os menores números que podem pertencer às diferentes classes (A Tabela 2 tem limites inferiores de classe de 50, 70, 90, 110 e 130).

Limites superiores de classe são os maiores números que podem pertencer às diferentes classes (A Tabela 2 tem limites superiores de classe de 69, 89, 109, 129 e 149).

Fronteiras de classe são os números usados para separar as classes, mas sem os saltos criados pelos limites de classe. A FIG. 1 mostra as lacunas criadas pelos limites de classe da Tabela 2 e, nessa figura, vemos que os valores de 69,5, 89,5, 109,5 e 129,5 estão no centro dessas lacunas e, seguindo o padrão dessas fronteiras de classe, vemos que a menor fronteira de classe é 49,5 e a maior é 149,5. Assim, a lista completa das fronteiras de classe é 49,5, 69,5, 89,5, 109,5, 129,5 e 149,5.

FIGURA 12 - ENCONTRANDO AS FRONTEIRAS DE CLASSE A PARTIR DOS LIMITES DE CLASSE EM UMA TABELA



Fonte: TRIOLA (2017).

Pontos médios das classes são os valores no centro da classe. A Tabela 2 tem pontos médios 59,5, 79,5, 99,5, 119,5 e 139,5. Cada ponto médio de classe pode ser encontrado somando-se o limite inferior de classe ao limite superior de classe e dividindo-se a soma por 2.

$$\text{Ponto Médio} = \frac{\text{limite inferior} + \text{limite superior}}{2}$$

Amplitude de classe é a diferença entre dois limites inferiores de classe consecutivos (ou duas fronteiras inferiores de classe consecutivas) em uma distribuição de frequência. A Tabela 2 usa uma amplitude de classe de 20.

Uma tabela de distribuição de frequência pode conter diferentes tipos de frequências, são elas:

Frequência Simples ou Frequência Absoluta da Classe (f_i): é o número de observações contadas dentro da classe.

Frequência Absoluta acumulada de Classe (F_i ou F_a): é a acumulação sucessiva, a partir da primeira classe até uma classe qualquer, das frequências simples ou absoluta das classes.

Frequência Relativa Absoluta de Classe (f_r): é a relação existente entre a frequência absoluta ou simples de classe e o número de observações da variável. Obtém-se a frequência relativa de cada classe a partir da seguinte equação: $f_r = \frac{f_i}{\sum f_i} \times 100$.

Importante!!! O símbolo \sum , é conhecido como somatório, ele representa a soma de todos os termos indicados à sua frente. No caso da frequência relativa absoluta de classe, o termo $\sum f_i$ significa que devemos somar todas as frequências simples da distribuição.

Frequência Relativa Acumulada (F_r): é a acumulação sucessiva, a partir da primeira classe até uma classe qualquer das frequências relativas das classes.

Para os dados da Tabela 2 temos os seguintes valores, correspondendo a cada frequência calculada, veja a Tabela 3.

TABELA 3 - MEDIDAS DE FREQUÊNCIA PARA OS ESCORES DE QI COMPLETOS DO GRUPO DE BAIXO NÍVEL DE CHUMBO

Score de QI	Frequência (f_i)	$f_r(\%)$	F_a	$F_r(\%)$
50 – 69	2	2,56	2	2,56
70 – 89	33	42,31	35	44,87
90 – 109	35	44,87	70	89,74
110 – 129	7	8,974	77	98,72
130 – 149	1	1,282	78	100,00
Total	78			

FONTE: ELABORAÇÃO PRÓPRIA (2023).

APRESENTAÇÃO DE DADOS DISCRETOS EM TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS

1. Se os dados são discretos, para organizar a tabela de distribuição de frequências: Conte quantas vezes cada dado se repete.
2. Organize a tabela, colocando, na primeira coluna, os dados numéricos em ordem natural (mas sem repetições) e, nas respectivas linhas, as frequências. Veja o seguinte exemplo: imagine que o gestor de um hospital veterinário quer apresentar o número de empregos que os candidatos à vaga recepcionista já tiveram. É preciso, então, contar quantas vezes cada valor se repete. Três candidatos nunca tiveram emprego

(número de empregos anteriores 0), dois candidatos tiveram um emprego e assim por diante. O gestor organiza os dados como mostra a Tabela 4.

TABELA 4 - DISTRIBUIÇÃO DOS CANDIDATOS À VAGA DE ESTÁGIO REMUNERADO EM UMA CLÍNICA VETERINÁRIA, SEGUNDO O NÚMERO DE EMPREGOS ANTERIORES.

Nº de empregos anteriores		Frequência
0		3
1		2
2		8
3		6
4		2
5		0
Total		21

FONTE: ELABORAÇÃO PRÓPRIA (2023).

APRESENTAÇÃO DE DADOS CATEGORIZADOS EM TABELA DE DISTRIBUIÇÃO DE FREQUÊNCIAS

Dados categorizados devem ser apresentados em uma tabela de distribuição de frequências, porque esse tipo de tabela pode mostrar quantas unidades foram observa-las em cada categoria da variável. A tabela 5 mostra dados categorizados, organizados em uma tabela de frequência para e homens e o número de mulheres que se candidataram à vaga de estágio remunerado em uma Clínica Veterinária.

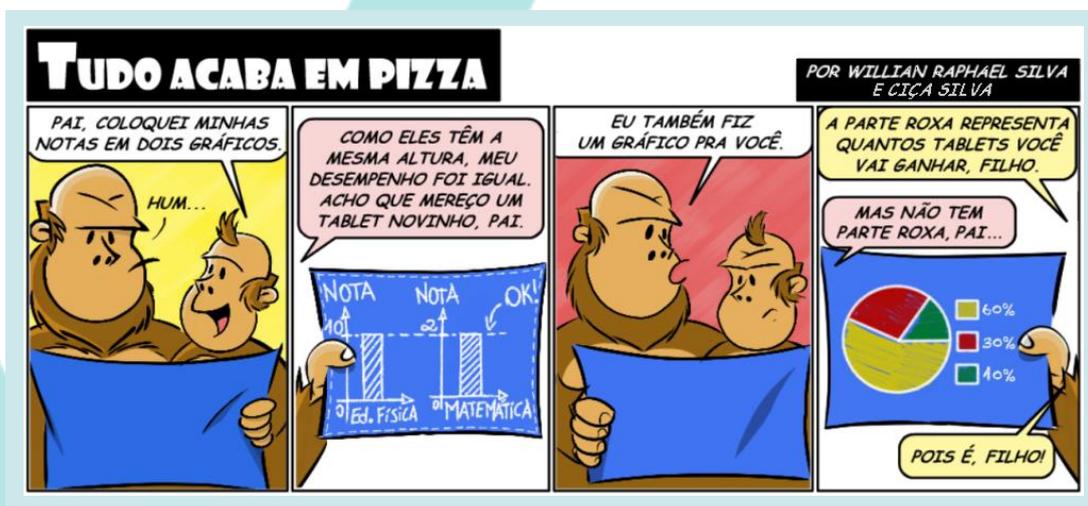
TABELA 5 - DISTRIBUIÇÃO DOS CANDIDATOS À VAGA DE ESTÁGIO REMUNERADO EM UMA CLÍNICA VETERINÁRIA, SEGUNDO O SEXO

Sexo	Frequência
Masculino	17
Feminino	15
Total	32

FONTE :ELABORAÇÃO PRÓPRIA (2023).

Representação por gráficos

FIGURA 13 - TUDO ACABA EM PIZZA



FONTE 13: [HTTPS://BLOGDOPROFH.WORDPRESS.COM/2014/10/15/NOTAS-E-GRAFICOS-TIRINHA/](https://blogdoprofh.wordpress.com/2014/10/15/notas-e-graficos-tirinha/) . ACESSO EM 16/01/2023.

Gráficos estatísticos são usados em jornais, em revistas e na televisão porque apresentam informação na forma de ilustrações, o que facilita a compreensão. Logo, é importante entender como os gráficos são construídos e o tipo de informação que podem fornecer.

APRESENTAÇÃO DE DADOS QUALITATIVOS

GRÁFICOS DE BARRAS: O gráfico de barras é um excelente recurso para apresentar dados qualitativos coletados em determinado momento, como as respostas de questionários.

FIGURA 15 - GRÁFICO DE BARRAS HORIZONTAL – FREQUÊNCIA DAS FRATURAS EXPOSTAS POR DIA DA SEMANA EM UM HOSPITAL



FONTE: GUIMARÃES, 2019. ADAPTADO

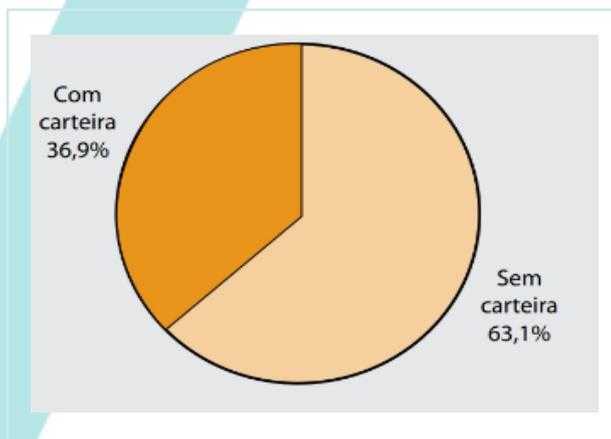
FIGURA 15 - GRÁFICO DE BARRAS VERTICAL – FREQUÊNCIA DAS FRATURAS EXPOSTAS POR DIA DA SEMANA EM UM HOSPITAL



FONTE: GUIMARÃES, 2019. ADAPTADO

Gráfico de setores: Os gráficos de setores são especialmente úteis para mostrar como se divide o todo. São popularmente conhecidos como gráficos de pizza, em razão do seu aspecto, divididos em fatias. Cada fatia é uma parte do todo.

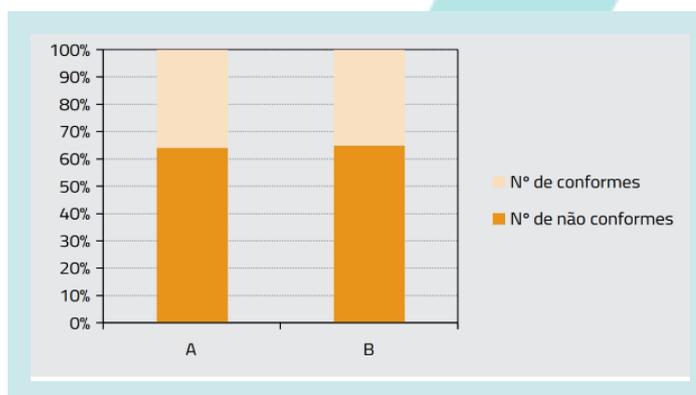
FIGURA 16 - DISTRIBUIÇÃO DOS PACIENTES DE UMA UBS CONFORME O PORTE DE CARTEIRA ASSINADA



FORNE: VIEIRA, SONIA, 2018, PÁG. 24

Gráfico Retangular de Composição: O gráfico retangular de composição ou gráfico de colunas empilhadas é uma alternativa para o gráfico de setores, porque também mostra o todo dividido em partes. É constituído por um retângulo com subdivisões, em que cada subdivisão representa uma categoria da variável. Esse tipo de gráfico deve ser usado para comparar dois grupos, que estejam subdivididos nas mesmas categorias.

FIGURA 17 - RESULTADO DO CONTROLE DE QUALIDADE DE MÁSCARAS CIRÚRGICAS



FORNE: VIEIRA (2018, P. 27).

Diagrama de Pareto: O diagrama de Pareto é um gráfico de barras ordenadas, das mais altas para as mais baixas. As categorias da variável ficam, portanto, ordenadas de acordo com as frequências, porque o comprimento da barra é dado pela frequência da categoria. O diagrama de Pareto é usado em gestão de qualidade quando se procuram os erros mais comuns, os motivos das perdas, as causas das reclamações.

FIGURA 18 - RECLAMAÇÕES NO PROCON POR ÁREA EM DETERMINADO ANO



FONTE 18: VIEIRA (2018, P. 28).

APRESENTAÇÃO DE DADOS QUALITATIVOS

Diagrama de pontos: Dados numéricos também podem ser apresentados em tabelas de distribuição de frequências. Se os dados são discretos, as tabelas de distribuição de frequências apresentam valores numéricos em ordem natural.

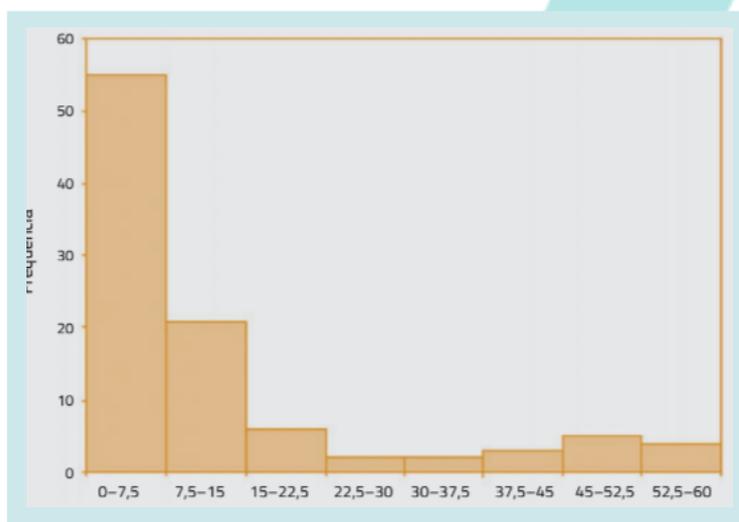
FIGURA 19 - AUSÊNCIAS DE PACIENTES NAS SESSÕES DE FISIOTERAPIA



FONTE: VIEIRA (2018, P. 30).

Histograma: Quando os dados são contínuos e a amostra é grande, não se deve fazer um gráfico de pontos. É mais conveniente condensar os dados, organizando uma tabela de distribuição de frequências e construir um histograma. O histograma é constituído por retângulos que têm base no intervalo de classe e área proporcional às frequências ou às frequências relativas (porcentagens) de cada classe. A área total sob o histograma é igual à soma das frequências, ou 100%.

FIGURA 20 - DISTRIBUIÇÃO DOS VALORES, EM REAIS, DEPOSITADOS EM CONTAS DE POUPANÇA QUE ANIVERSARIARAM NO DIA 1º DE JUNHO DE DETERMINADO ANO.



FONTE: VIEIRA (2018, P. 31).

Diagrama de Ramo e folhas: Se você tiver dados contínuos, em quantidade não muito grande (digamos entre 25 e 100), pode fazer o diagrama de ramo e folhas, que mostra bem a forma da distribuição. O diagrama é constituído de duas partes, o ramo e as folhas, nos quais os

números estão ordenados. Os primeiros dígitos dos dados numéricos são os ramos e os últimos dígitos são as folhas.

FIGURA 21 - DIAGRAMA DE RAMO E FOLHAS PARA A DISTRIBUIÇÃO DAS IDADES DOS FUNCIONÁRIOS DE UMA CLÍNICA

1	8	9	9						
2	0	1	2	2	4	5	6	7	9
3	0	0	0	1	1	2	6	7	
4	1	2	6						
5	1	9							

FONTE: VIEIRA (2018, P. 32).



LEITURA COMPLEMENTAR

Acesse o link abaixo e veja passo a passo como construir os gráficos que acabamos de estudar.

<https://integrada.minhabiblioteca.com.br/reader/books/9788522128082/pageid/32>

MEDIDAS DE POSIÇÃO OU TENDÊNCIA CENTRAL

Para entender as características gerais de um conjunto de dados, as pessoas preferem olhar uma figura. Como diz o ditado popular, “uma imagem vale por mil palavras”. Daí a importância dos métodos gráficos descritos na seção anterior. Muitas vezes, porém, é preciso fornecer um resumo dos dados. Isso pode ser feito por meio das medidas de tendência central, que dão ideia do centro em torno do qual os dados se distribuem.

Para descrever a tendência central dos dados, calculamos **média**, **mediana** e **moda**.

MÉDIA

A média aritmética, ou simplesmente média, é a medida de tendência central mais conhecida e utilizada para resumir a informação contida em um conjunto de dados. Vamos estudar a média para dados simples e para dados apresentados em tabelas de distribuição de frequências, o que difere é o modelo matemático que utilizamos para determinar o valor da média. Vejamos:

- **Dados Simples** - A média é obtida somando todos os dados e dividindo o resultado pelo número deles.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + x_4 + \dots + x_n}{n} = \frac{\sum x_i}{n}$$

Dados observados

Número de observações



INDICAÇÃO DE VÍDEO

Clique no link abaixo e veja passo a passo de como calcular a média para dados simples.

<https://youtu.be/vJ0A36fs5Gs>

- **Dados em tabela de distribuição de frequências** – Neste caso temos os dados contínuos e dados discretos, vejamos como se calcula a média em cada caso:
- **Dados discretos em tabela de frequências:** a média de dados discretos agrupados em uma tabela de distribuição de frequências é dada pela soma dos produtos dos valores da variável pelas respectivas frequências, dividida pela soma das frequências.

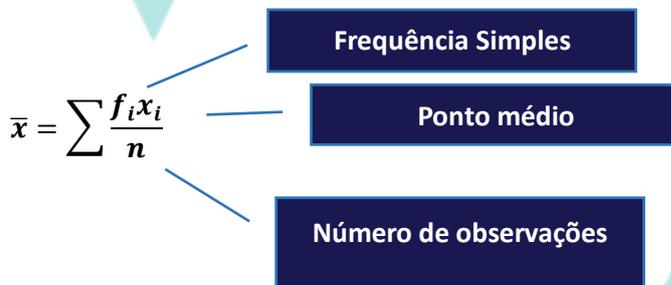


INDICAÇÃO DE VÍDEO

Clique no link abaixo e veja passo a passo calcular a média para dados discretos.

<https://youtu.be/8he1mp3RPU4>

- **Dados contínuos em tabela de frequências:** dados contínuos podem estar agrupados em classes e ser apresentados em tabelas de distribuição de frequências. Nesses casos, antes de calcular a média, é preciso calcular o ponto médio de cada classe. A fórmula matemática é semelhante à de dados discretos, porém o símbolo x_i aqui corresponde ao ponto médio de cada classe.



INDICAÇÃO DE VÍDEO

Clique no link abaixo e veja passo a passo calcular a média para dados contínuos.

<https://youtu.be/0gLFm8LySzY>

MEDIANA

Mediana é o valor que ocupa a posição central do conjunto dos dados ordenados.

Se o número de dados é ímpar, existe um único valor na posição central. Esse valor é a mediana. Por exemplo, dados: 3, 5 e 9, a mediana é 5.

Se o número de dados é par, existem dois valores na posição central. Então, a mediana é a média desses dois valores centrais. Dados: 3, 5, 7 e 9, a mediana é 6, isto é, a média de 5 e 7.

Quando os dados estão em tabela de distribuição de frequências a mediana pode ser calculada das seguintes formas:

- **Dados discretos em tabela de frequências:** é calculada a posição mediana (P_{md}) por meio da fórmula $P_{md} = \frac{\sum f_i}{2}$, em seguida através da frequência acumulada encontra-se a classe que contém a posição mediana, o valor correspondente à classe é a mediana.

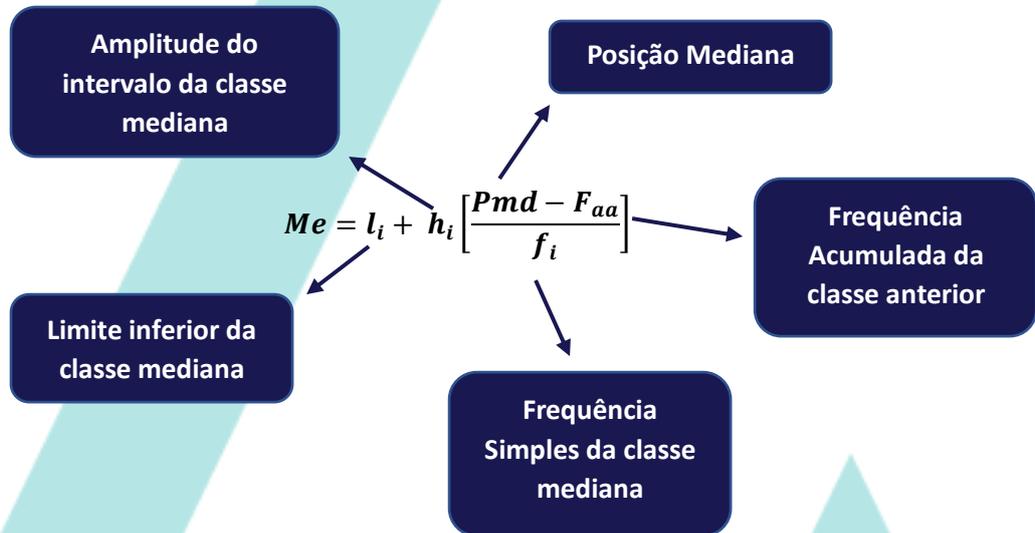


INDICAÇÃO DE VÍDEO

Clique no link abaixo e veja passo a passo calcular a mediana para dados discretos.

<https://youtu.be/jEf0ZtSuYEQ>

- **Dados contínuos em tabela de frequências:** neste caso, usamos a seguinte fórmula matemática:



INDICAÇÃO DE VÍDEO

Clique no link abaixo e veja passo a passo calcular a mediana para dados contínuos.

<https://youtu.be/FOPZyKkPH4I>

MODA

Moda é o valor que ocorre com maior frequência.

Um conjunto de dados pode não ter moda, ou ter duas ou mais modas. O conjunto de dados

0, 2, 4, 6, 8, 10

não tem moda e o conjunto

1, 2, 2, 3, 4, 4, 5, 6, 7

tem duas modas: 2 e 4. A moda de uma amostra é, em geral, indicada por **mo**.

A moda também pode ser usada para descrever dados qualitativos. Nesse caso, a moda é a categoria que ocorre com maior frequência, ou seja, a categoria que concentra maior quantidade de dados.

Quando os dados estão em tabela de distribuição de frequências, a mediana pode ser calculada das seguintes formas:

- **Dados discretos em tabela de frequências:** a moda é a medida ou informação que possui maior frequência absoluta.

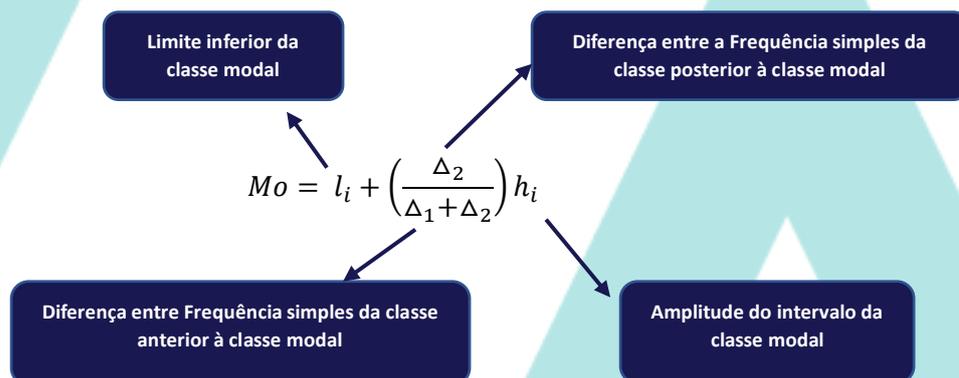


INDICAÇÃO DE VÍDEO

Clique no link abaixo e veja passo a passo calcular a moda para dados discretos.

<https://youtu.be/DDI1AUuurCE>

- **Dados contínuos em tabela de frequências:** neste caso, usamos a seguinte fórmula matemática:



INDICAÇÃO DE VÍDEO

Clique no link abaixo e veja passo a passo calcular a moda para dados contínuos.

<https://youtu.be/LrXb3yJgMPk>

MEDIDAS DE DISPERSÃO/VARIABILIDADE

Mínimo, máximo e amplitude

- ✓ Mínimo de um conjunto de dados é o número de menor valor.
- ✓ Máximo de um conjunto de dados é o número de maior valor.
- ✓ Amplitude de um conjunto de dados é a diferença entre o valor máximo e o valor mínimo do conjunto, ou seja, $amplitude = máximo - mínimo$

DESVIO PADRÃO

O desvio padrão é uma medida de variabilidade muito recomendada, porque mede de maneira eficaz a dispersão dos dados em torno da média. Mas, para calcular o desvio padrão, é preciso, primeiro, calcular a variância. Vamos, então, entender o que é variância.

- **Variância**

Quando a média é usada como medida de tendência central, ou seja, quando a média indica o centro da distribuição, podemos calcular o desvio de cada observação em relação à média como se segue:

$$desvio = observação - média \quad desvio = x - \bar{x}$$

Se os desvios em relação à média são pequenos, podemos concluir que as observações estão aglomeradas em torno da média. A variabilidade dos dados é, portanto, pequena.

Se os desvios são grandes, os dados estão muito dispersos, ou seja, a variabilidade é grande. A variância é uma medida de variabilidade que capta essas situações.

A variância é determinada pelo modelo matemático:

	Amostra		População	
Dados simples	$s^2 = \frac{\sum(x_i - \bar{x})^2}{n - 1}$	$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$	$s^2 = \frac{\sum(x_i - \bar{x})^2}{n}$	$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n}$
Dados Discretos ou contínuos	$s^2 = \frac{\sum f_i(x_i - \bar{x})^2}{n - 1}$		$s^2 = \frac{\sum f_i(x_i - \bar{x})^2}{n}$	

Agora que vimos como determinamos a variância, podemos continuar nossos estudos sobre o desvio padrão.

O desvio padrão é a raiz quadrada da variância, com sinal positivo. Sendo assim, temos:

$$s = \sqrt{\text{variância}}$$

COEFICIENTE DE VARIAÇÃO

Coeficiente de variação é a razão entre o desvio padrão e a média. O coeficiente de variação não tem unidade de medida, ou seja, é adimensional. Isso acontece porque média e desvio padrão são medidos na mesma unidade, que então se cancelam. Pode, portanto, ser expresso em porcentagem. Então:

$$CV = \frac{s}{\bar{x}} \times 100$$

O coeficiente de variação mede a dispersão dos dados em relação à média. Por ser adimensional, é útil para comparar a dispersão relativa de variáveis medidas em diferentes unidades.

MEDIDAS DE POSIÇÃO RELATIVA E DIAGRAMA DE CAIXA

Esta seção introduz medidas de posição relativa, que são números que mostram a localização de valores de dados em relação a outros valores dentro do mesmo conjunto de dados. O conceito mais importante é o de escore z, o qual será usado frequentemente nas unidades a seguir. Discutimos, também, percentis e quartis, que são estatísticas comuns, bem

como um novo gráfico estatístico, chamado diagrama em caixa.

ESCORES Z

Encontra-se um escore z (ou valor padronizado) convertendo-se um valor para uma escala padronizada, como dado na definição a seguir. Essa definição mostra que um escore z é o número de desvios-padrão que um valor de dado está afastado da média (TRIOLA, 2017).

DEFINIÇÃO

Um escore z (ou valor padronizado) é o número de desvios-padrão que um dado valor x está acima ou abaixo da média. O escore z é calculado usando-se uma das seguintes fórmulas:

$$z = \frac{x - \bar{x}}{s}$$

Para os valores de escores z, arredonde-os para duas casas decimais, isto é, motivado pelo formato das tabelas-padrão nas quais os escores z são expressos com duas casas decimais.



INDICAÇÃO DE VÍDEO

Agora entenda como aplicamos esse conceito na prática.

<https://youtu.be/AhMCOWvpwVc>

PERCENTIL

Percentis são um tipo de quantis que fazem a partição dos dados em grupos com, basicamente, o mesmo número de valores em cada grupo.

DEFINIÇÃO

Percentis são medidas de localização, denotados por P1, P2, ..., P99, que dividem um conjunto de dados em 100 grupos com cerca de 10% dos valores em cada um deles.

O processo para se encontrar o percentil que corresponde a determinado valor x é dado pela expressão seguinte:

$$\text{Percentil do valor } x = \frac{\text{n}^\circ \text{ de valores menores do que } x}{\text{n}^\circ \text{ total de valores}} \cdot 100$$



INDICAÇÃO DE VÍDEO

Veja aqui um exemplo de cálculo de percentil:

<https://youtu.be/4AuBPtFyGhc>

QUARTIS

Assim como há 99 percentis que dividem os dados em 100 grupos, há três quartis que dividem os dados em quatro grupos.

DEFINIÇÃO

Quartis são medidas de localização, denotadas por Q_1 , Q_2 , Q_3 , que dividem um conjunto de dados em quatro grupos, com cerca de 25% dos valores em cada grupo.

Eis uma descrição dos três quartis mais precisa que a apresentada na definição anterior:

Q₁ (Primeiro quartil)	Mesmo valor de P ₂₅ . Separa os 25% inferiores dos 75% superiores dos valores ordenados. (Mais precisamente, pelo menos 25% dos valores ordenados são menores do que ou iguais a Q ₁ e, no mínimo, 75% dos valores são maiores do que ou iguais a Q ₁ .)
Q₂ (Segundo quartil)	Mesmo que P ₅₀ e mesmo que a mediana; separa os 50% valores ordenados inferiores dos 50% superiores.
Q₃ (Terceiro quartil)	Mesmo que P ₇₅ . Separa os 75% valores ordenados inferiores dos 25% superiores. (Mais precisamente, no mínimo 75% dos valores ordenados são menores do que ou iguais a Q ₃ , e no mínimo 25% dos valores são maiores do que ou iguais a Q ₃ .)

$Q_1 = P_{25}$; $Q_2 = P_{50}$ e $Q_3 = P_{75}$



INDICAÇÃO DE VÍDEO

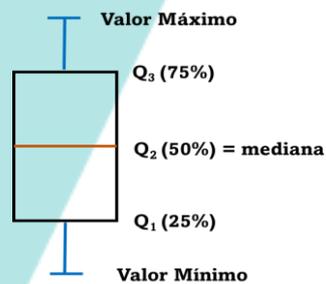
Quer entender melhor? Então não deixe de assistir ao vídeo

<https://youtu.be/ie2few5B7ZU>

DIAGRAMA DE CAIXA

Um diagrama em caixa ou boxplot (ou diagrama em caixa e bigode) é um gráfico de um conjunto de dados que consiste em uma linha a qual se estende do valor mínimo ao valor máximo, e em uma caixa com linhas traçadas no primeiro quartil, Q₁, na mediana e no terceiro quartil, Q₃. A FIG. 10 mostra a aparência de um boxplot.

FIGURA 22 - ELEMENTOS DE UM GRÁFICO BOXPLOT



FONTE: ELABORAÇÃO PRÓPRIA (2023).



INDICAÇÃO DE VÍDEO

Para saber mais e aprender como construir um diagrama de caixa na prática, assita ao vídeo a seguir:

<https://youtu.be/ebkzrOMjGXU>



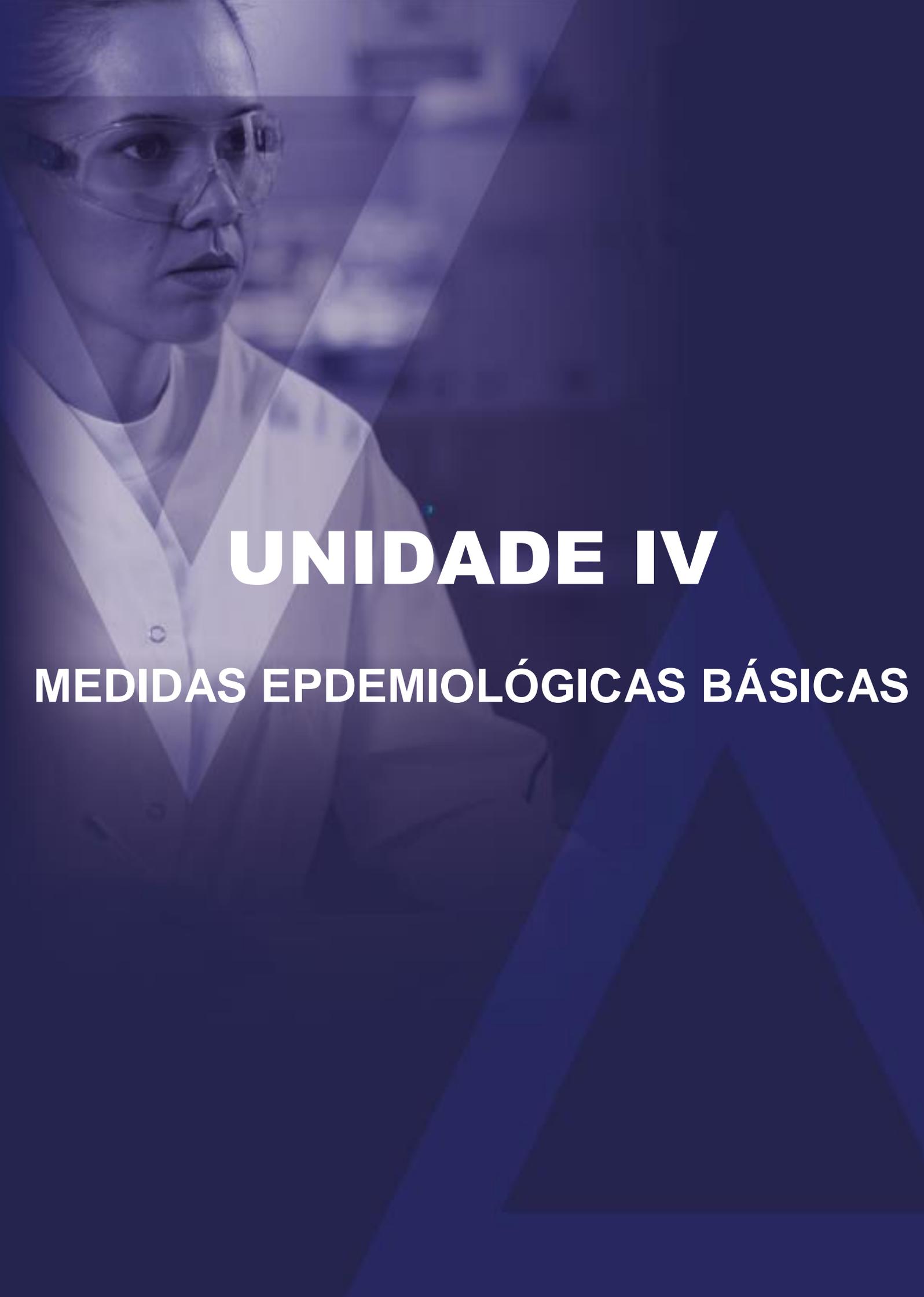
RESUMO DA UNIDADE

Nesta unidade, abordamos as principais medidas descritivas que são necessárias para uma análise dos dados de uma pesquisa ou estudo. Dados corretamente coletados fornecem conhecimentos que não seriam obtidos por simples especulação. Nas mais diversas áreas de trabalho, as pessoas levantam dados e os analisam para tirar conclusões, mas os dados precisam ser apresentados. A melhor forma de apresentar dados de maneira organizada é por meio de tabelas e gráficos que devem ser elaborados de acordo com o tipo de dado que sua pesquisa investiga.

Mesmo que a informação visual fornecida pelos gráficos e tabelas sejam atraentes e, muitas vezes, a maior preferência, é preciso fornecer um resumo dos dados. Isso pode ser feito por meio das medidas de tendência central, que dão ideia do centro em torno do qual os dados se distribuem.

A variabilidade é inerente aos fenômenos físicos, naturais e econômicos. Por exemplo, na mesma cidade a temperatura varia ao longo do dia e, no mesmo dia, registram-se temperaturas muito diferentes em diferentes lugares do mundo. O peso das pessoas varia ao longo da vida. Assim, média, mediana e moda não bastam para descrever um conjunto de dados, elas indicam o centro em torno do qual os dados estão dispersos, mas não revelam o quanto se dispersam. Para descrever um conjunto de dados, você deve apresentar, além da medida de tendência central, uma medida de variabilidade ou dispersão. Por isso, estudamos as medidas usadas para medir variabilidade: mínimo e máximo, amplitude, quartis, percentis, variância, desvio padrão.

Dessa forma, esta unidade explorou por completo o universo da análise descritiva de um conjunto



UNIDADE IV

MEDIDAS EPIDEMIOLÓGICAS BÁSICAS



OBJETIVOS

- ✓ Compreender as principais medidas epidemiológicas de morbidade e mortalidade, formas de cálculo e interpretações.



META DA UNIDADE

A análise da situação de saúde é um dos usos da Epidemiologia na interface entre a produção de conhecimentos e sua aplicação aos serviços de saúde. Isso significa conhecer o que afeta a saúde da população e em que medida isso ocorre; o que compromete o desempenho das potencialidades humanas; o que produz doenças; o que provoca óbitos, sobretudo os que se podem evitar com prevenção e assistência adequada. Em especial, interessa-nos saber como as doenças e agravos de interesse em saúde coletiva se distribuem segundo as condições devida.

Os indicadores epidemiológicos de morbi-mortalidade são ferramentas que constituem a estratégia descritiva, que busca estudar as características e do impacto da doença sobre a situação de saúde na população. Essa estratégia, aplicada ao estudo de uma doença ou agravo à saúde, inclui o uso de métodos epidemiológicos que revelem os padrões de distribuição no tempo, no espaço geográfico e de acordo com atributos dos indivíduos acometidos e das coletividades humanas.

Nesta unidade, estudaremos os indicadores epidemiológicos de morbidade, que são a prevalência e a incidência; e os indicadores de mortalidade, envolvendo as taxas de mortalidade geral e específica, com destaque para os coeficientes de mortalidade infantil e materna e a letalidade.

Afinal de contas, o que são indicadores epidemiológicos?

Os indicadores epidemiológicos são parâmetros que servem para quantificar ou medir a frequência com que os problemas de saúde ocorrem nas populações e, por isso, representam a ligação entre casos de doenças ou de óbitos para a população, considerando um intervalo de tempo. Os indicadores são utilizados para fazer estimativas sobre o risco de adoecimento em decorrência de agravo da doença na população para um período de tempo.

EXEMPLO

O número de profissionais de saúde, de unidades básicas de saúde, de leitos hospitalares, de consultas e exames disponíveis para população estudada são indicadores que mensuram os recursos em saúde disponíveis para atendimento da população. Isso porque permite compreender se as más condições de saúde de uma população seriam resultado de contingente e infraestrutura insuficientes ou se esses estariam proporcionais à população estudada, porém, a gestão deveria ser repensada para melhorar os fluxos de trabalho e, assim, implantar as estratégias de modo efetivo

MEDIDAS DE MORBIDADE

As estatísticas de morbidade têm como característica fundamental o fato de serem utilizadas, preferencialmente, para avaliação do nível de saúde (vida) e para o aconselhamento de ações de caráter abrangente que visem melhorar o estado sanitário (estrada) da comunidade (saneamento básico, por exemplo). Para aproximar a correção das decisões ao apoiar ações específicas necessárias ao controle de determinada doença (escolha de um tratamento para hanseníase, por exemplo), consultam-se os índices de morbidade discriminados em coeficientes de incidência e de prevalência específicos para as múltiplas variáveis independentes (MARTINS, 2018).

Sob o ponto de vista geral, denominam-se índices de morbidade as medidas relativas ao comportamento das doenças e dos agravos à vida em uma população exposta.

A expressão quantitativa da morbidade é dada por diferentes índices de morbidade, dentre os quais se destacam os coeficientes de morbidade gerados para fins operacionais, definidos pelos quocientes entre o número de casos de uma doença e o número de indivíduos da população. Através desse modelo matemático, deixam de sofrer a influência do tamanho da população e passam a expressar os riscos de ocorrência e as chances de existência da morbidade em uma população, comparáveis aos de outras populações, de outros temp

$$\text{Coeficiente de morbidade} = \frac{\text{Número de casos de uma doença}}{\text{população}} \cdot 10^n$$

PREVALÊNCIA

Em epidemiologia é a frequência de casos de morbidade ou outra condição qualquer que se destaca na população. É entendida como a proporção de uma população que tem a doença em determinado momento. Assim, em saúde pública, prevalência é termo descritivo da força com que subsistem as doenças nas comunidades (MARTINS, 2018).

A taxa de prevalência permite realizar comparações e análises do quadro sanitário de diferentes populações. Define-se prevalência como a relação entre o número de casos conhecidos de uma dada doença e a população de origem dos casos, com referência a um lugar definido, multiplicando-se o resultado pela base referencial da população que é potência de 10, em geral 100, ou por 1.000, 10.000, 100.000.

$$\text{Prevalência} = \frac{\text{Número de casos conhecidos de uma doença}}{\text{população}} \cdot 10^n$$

Em outras palavras, a prevalência representa o estoque de casos, isto é, a proporção da população que apresenta uma dada doença. Ela aumenta com os casos novos e decresce com o número de curas e de óbitos. A FIG. 1 é uma ilustração de como se processa a prevalência.

FIGURA 23 - FATORES QUE INFLUENCIA NA TAXA DE PREVALÊNCIA.



FONTE: [HTTPS://2.BP.BLOGSPOT.COM/-RHJDJ9BTWYC/WCO-HOFC9HI/AAAAAAAAAALk/WTEkP5NmVYcNoRUxGM_HkyQANTRI7XZxACLcB/s1600/PREVALENCIA.JPG](https://2.bp.blogspot.com/-RHJDJ9BTWYC/WCO-HOfc9HI/AAAAAAAAAALk/WTEkP5NmVYcNoRUxGM_HkyQANTRI7XZxACLcB/s1600/PREVALENCIA.JPG). ACESSO EM 17/01/2023.

A prevalência é muito útil em planejamento e administração de serviços e de programas. O conhecimento da prevalência pode ser o mais indicado quando se pretende colocar à disposição da população um determinado serviço de saúde ou produto, como é o caso de um programa para tratamento antiparasitário em massa ou para fornecer óculos a escolares com deficiências visuais: no caso, é necessário saber a prevalência da parasitose ou de escolares deficientes visuais não-possuidores de óculos. Quando se trata de prever necessidade de serviços, tais como número de leitos, consultas e pessoal, o conhecimento da prevalência é, na maioria das vezes, suficiente (GORDIS, 2017).

INCIDÊNCIA

A incidência diz respeito à frequência com que surgem novos casos de uma doença em um intervalo de tempo, como se fosse um “filme” sobre a ocorrência da doença, no qual cada quadro pode conter um novo caso ou novos casos (PEREIRA, 1995). É, assim, uma medida dinâmica.

$$\text{Incidência} = \frac{\text{N}^\circ \text{ de casos novos de uma doença ocorrentes em determinada comunidade em certo período de tempo}}{\text{N}^\circ \text{ de pessoas expostas ao risco de adquirir a doença no referido período}} \cdot 10^n$$

A incidência é a medida mais importante da epidemiologia. Ela é a preferida em investigações científicas, seja nas pesquisas etiológicas, em estudos de prognósticos, na verificação da eficácia das ações terapêuticas e preventivas e em outros tipos de pesquisa. Seu conhecimento, ou estimativa aproximada, é necessário para planejar as investigações - como os ensaios clínicos - e, em especial, para determinar o tamanho da amostra (MARTINS, 2018).

Em muitas situações, no entanto, o conhecimento da incidência, embora desejável, não é de obtenção simples, como no caso nas doenças crônicas. Na impossibilidade de medir diretamente a incidência, usa-se a prevalência, cuidadosamente, como substituta da incidência, ou estima-se a incidência a partir de dados de prevalência.

MEDIDAS DE MORTALIDADE

Historicamente, o primeiro indicador utilizado em avaliações de saúde coletiva, e ainda hoje o mais empregado, é o de mortalidade. Isso pode ser explicado pelas facilidades operacionais: a morte é objetivamente definida, ao contrário da doença, e cada óbito tem de ser registrado. O registro obrigatório resulta na formação de uma “base de dados”, mantida e atualizada por técnicos do governo, divulgada periodicamente. Essa base de dados é armazenada em computador, o que facilita a preparação de estatísticas sob diversas formas, cuja interpretação, mesmo que superficial, fornece um diagnóstico da situação (MARTINS, 2018). Medidas de mortalidade têm como utilização preferencial avaliação do nível de saúde e indicação de medidas preventivas e de controle de caráter abrangente (saneamento básico, detecção precoce do câncer de mama, redução do tabagismo, por exemplo), que objetivem melhorar o estado sanitário da comunidade. Essa aplicação visa, por uma parte, sugerir ações que reduzam o risco de morrer por uma determinada causa evitável e, por outra parte, indicar a necessidade de ações de controle sobre fatores de risco de adoecer ou de sofrer agravos que se associem a alta letalidade. A avaliação da efetividade dessas ações poderá ser feita pelo acompanhamento da evolução das medidas de mortalidade (GORDIS, 2017).

INDICADORES DE MORTALIDADE

A taxa de mortalidade é definida, de forma mais geral, como a proporção entre a frequência absoluta de óbitos e o número de indivíduos expostos ao risco de morrer (população exposta), no mesmo período de referência e no mesmo local.

$$\text{Taxa de Mortalidade} = \frac{\text{N}^{\circ} \text{ de óbitos}}{\text{População exposta}} \cdot 10^n$$

A depender de como se define a população considerada exposta ao risco de óbito em determinado lugar e período de tempo, pode-se calcular a taxa de mortalidade geral e os indicadores de mortalidade específicos. Esses últimos são relativos a grupos populacionais definidos de acordo com categorias de: sexo, idade, raça/etnicidade, causa do óbito, grupo socioeconômico ou outra característica de interesse (ROUQUAYROL, 2017).

TAXA DE MORTALIDADE GERAL

A taxa de mortalidade geral (TMG) é calculada dividindo-se o número total de óbitos por todas as causas, em um determinado ano, pela população daquele ano, registrados em uma determinada área e multiplicando-se por 1.000, base referencial para a população exposta.

$$\text{Taxa de Mortalidade Geral} = \frac{\text{N}^{\circ} \text{ de todos óbitos}}{\text{População}} \cdot 10^n$$

É conhecido como taxa bruta de mortalidade quando seus valores não são padronizados. A TMG tem importante aplicação na avaliação do estado sanitário de populações e áreas definidas, associado a outros coeficientes, taxas e índices. Esse indicador possibilita relacionar o nível de saúde de regiões diferentes em uma mesma época ou de se empreenderem estudos comparativos, nos quais seja considerada a variável tempo. Apesar de ser um dos indicadores mais utilizados em epidemiologia, na prática a sua aplicação em estudos comparativos sofre limitações em decorrência de variações na organização dos serviços de saúde, na qualidade dos registros de óbitos entre áreas e por diferenças na estrutura etária das populações comparadas (ROUQUAYROL, 2017).

TAXA DE MORTALIDADE AJUSTADA

Sendo a TMG um indicador que não leva em conta fatores específicos da população estudada como faixa etária, sexo, etnia e classe social, entre outros, não é o coeficiente mais apropriado para comparação dos riscos de mortalidade de diferentes populações. Para isso, devemos utilizar a taxa de mortalidade ajustada ou padronizada, que é calculada através de técnicas de padronização direta ou indireta que garantem o uso de uma população padrão como referência. Quando são realizadas comparações entre os riscos de morte de diferentes países, a população padrão é a mundial. Para as comparações nacionais de regiões do mesmo país, a população de escolha padrão é a do próprio país (MARTINS, 2018).

MÉTODO DE PADRONIZAÇÃO DIRETA

A utilização desse método exige que se disponha do total de óbitos, distribuídos por grupos do fator em estudo: por exemplo, se o fator de estudo é a faixa etária, é necessário conhecer o total de óbitos, distribuídos por grupos de idade, e das distribuições etárias das

populações em estudo. A partir dessas informações é possível calcular os coeficientes específicos de cada faixa etária. Aplicando-se esses coeficientes a uma distribuição etária padrão, será calculado o total de óbitos, supondo que a estrutura etária dos países ou regiões de interesse sejam semelhantes à padrão. O quociente entre o total de óbitos estimados e a população padrão é a TGM padronizada. Resumidamente, a padronização direta pode ser obtida através da seguinte fórmula matemática:

$$PD = \frac{(\sum \text{Taxas específicas}) \cdot (\text{população padrão})}{\text{população de estudo}}$$

$$\text{taxa específica} = \frac{\text{N}^\circ \text{ de óbitos por determinada doença}}{\text{População}} \cdot 10^n$$

MÉTODO DE PADRONIZAÇÃO INDIRETA

Esse método é empregado quando não está disponível a informações necessárias para o emprego do método de padronização direta. Esse método consiste na estimação da função de mortalidade da população de cujos dados necessários através da utilização dos coeficientes específicos de uma população padrão escolhida não dispomos. Nesse caso, assumimos que a população padrão escolhida em uma função com a mesma forma ou estrutura da(s) população(ões) em estudo, independentemente dos níveis semelhantes ou não (ROUQUAYROL, 2017).

Para que a padronização indireta seja possível, é necessário que se tenha o total de óbitos e a distribuição etária das populações em estudo. Em linhas gerais, esse método consiste na aplicação dos coeficientes específicos da população padrão às populações em estudo, estimando-se, assim, o número de óbitos esperados, caso as populações tivessem os mesmos coeficientes que a população-padrão. Feito isso, obtém-se a constante SMR (*standardized mortality ratio*),

$$SMR = \frac{\text{total de óbitos observados}}{\text{total de óbitos esperados}} \cdot \text{coeficientes específicos da população padrão}$$

E assim obtemos a padronização indireta por:

$$PI = \text{Taxa População Padrão} \cdot SMR$$

O maior desafio desse método é a escolha de uma função padrão que apresente a mesma forma da população em estudo. Dependendo da população escolhida, os resultados do processo de padronização indireta podem sub ou superestimar as diferenças entre os níveis de mortalidade da(s) população(ões) de interesse (GORDIS, 2017).

TAXA DE MORTALIDADE INFANTIL

O taxa de mortalidade infantil (TMI) é calculada pela divisão do número de óbitos de crianças menores de um ano pelos nascidos vivos naquele ano, em determinada área, multiplicando por 1.000 o valor encontrado. Esse coeficiente mede, portanto, o risco de morte para crianças menores de um ano. O CMI possui dois componentes: mortalidade neonatal e mortalidade pós-neonatal. A mortalidade neonatal inclui os óbitos infantis com menos de 28 dias do nascimento, enquanto a mortalidade pós-neonatal ou tardia compreende os óbitos entre 28 dias e um ano de vida. O coeficiente de mortalidade neonatal pode ser considerado indicador sensível de subdesenvolvimento, pois, quando é elevado, aponta para falhas nos sistemas de proteção e promoção da saúde infantil (ROUQUAYROL, 2017). Os dados absolutos, que permitem o cálculo desse coeficiente, podem ser influenciados pela qualidade dos registros, tanto de nascidos vivos quanto de óbitos de menores de um ano. Trata-se de um dos indicadores epidemiológicos mais sujeitos a distorções, dentre elas, ausências de registros de nascimentos e de óbitos, definição de nascido vivo no ano, declarações erradas da causa de morte e da idade da criança. Apesar das distorções, ainda assim o CMI é um indicador razoável de desigualdades regionais em saúde.

$$TMI = \frac{\text{N}^\circ \text{ de óbitos de crianças } < 1 \text{ ano}}{\text{nascidos vivos}} \cdot 1000$$

COMENTÁRIO DO AUTOR

Indicadores como taxa de mortalidade infantil, taxa de mortalidade por doenças transmissíveis, esperança de vida ao nascer, razão de mortalidade proporcional e coeficiente geral de mortalidade são exemplos de indicadores que permitem enxergar o quadro de saúde dos grupos analisados. Isso porque a sucessiva coleta de dados desses indicadores ao longo dos anos possibilita traçar uma linha do tempo em que se pode observar a evolução dos parâmetros de saúde. A partir do entendimento dos resultados apresentados nos indicadores, os órgãos de saúde pública, partindo desse panorama, conseguem apontamentos que permitem rastrear os fatores que conduziram ao perfil de saúde apresentado nos resultados (ROUQUAYROL, 2017).

TAXA DE LETALIDADE

Letalidade expressa o maior ou menor poder que tem uma doença ou agravo de provocar a morte das pessoas acometidas pela doença ou que sofreram agravos à saúde, ou seja, o risco de morte por uma doença é medido a partir da taxa de letalidade. Para estimar a medida dessa propriedade da morbidade, calcula-se o coeficiente de letalidade, dado por:

$$TL = \frac{N^{\circ} \text{ óbitos devidos a determinada causa}}{N^{\circ} \text{ de pessoas que foram realmente afetadas pela doença}} \cdot 100$$

A letalidade das doenças varia com a idade, o sexo, as condições socioeconômicas, o estado imunitário do indivíduo, a virulência do bioagente, nos casos das doenças infecciosas, e a eficácia do tratamento. Assim, existem doenças cuja letalidade esperada era de 100%, como é o caso da raiva humana até anos recentes, quando todos os pacientes que desenvolviam sinais e sintomas da doença morriam. Também existem doenças cuja letalidade esperada é zero, como é o caso do resfriado comum. (MARTINS, 2017)

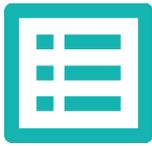
De maneira geral, a partir de evidências científicas ou de dados estatísticos é possível estimar o risco que um indivíduo tem de morrer a partir de uma doença. Como exemplo, tem-se a letalidade por leptospirose, cujo coeficiente médio geral esperado é de 10%, ou seja, de cada 100 pacientes que desenvolvem a doença, 10 poderão morrer em consequência dela. O coeficiente de letalidade de uma doença possibilita, ainda que de maneira indireta, avaliar a resolutividade dos serviços de saúde e o acesso a estes. Também podem ser citados como

exemplos os agravos por doença meningocócica e dengue, cujos diagnósticos precoces e a qualidade da assistência contribuem para a redução da letalidade. (ROUQUAYROL, 2017)



LEITURA COMPLEMENTAR

MERCHAN-HAMANN, E.; TAUIL, P. L.; COSTA, M. P. Terminologia das medidas e indicadores em epidemiologia: subsídios para uma possível padronização da nomenclatura. Informe Epidemiológico do Sus, Brasília, DF, v. 9, n. 4, p. 276-284, dez. 2000. Disponível em: http://scielo.iec.gov.br/scielo.php?script=sci_arttext&pid=S0104-1673200000040006 >. Acesso em: 23 jan. 2023.



RESUMO DA UNIDADE

Nesta unidade discutimos os principais indicadores empregados na epidemiologia para abordar o estado de saúde da população.

Indicador de saúde se refere às medidas que sintetizam informações sobre características da situação de saúde, bem como aponta a efetividade do sistema de saúde. Quando associados, os indicadores apontam as condições sanitárias da população estudada e auxiliam na vigilância da qualidade da saúde. Assim, os indicadores constituem uma via de avaliação da condição de saúde atual, permitem avaliar a evolução da saúde das populações ao longo do tempo e traçar diversos comparativos.

A morbidade se refere ao conjunto de indivíduos que adquiriram determinada doença em um intervalo de tempo.

A incidência de uma doença expressa a dinâmica com que os casos aparecem no grupo. Por exemplo, ela informa quantos, entre os sadios, tornam-se doentes em dado período de tempo, ou então quantos, entre os doentes, apresentam complicação ou morrem, decorrido algum tempo. Usualmente, diz-se que a incidência reflete “a força da morbidade” (ou “força da mortalidade”, em se tratando de óbitos). Portanto, a incidência de doenças, em uma população, significa a ocorrência de novos casos relacionados à unidade de intervalo de tempo, seja ele dia, semana, mês ou ano.

A prevalência descreve a força com que subsistem as doenças na coletividade. Portanto, trata-se de um indicador de morbidade. A taxa de prevalência, por sua vez, indica uma medida que permite estimar e comparar, no tempo e no espaço, a ocorrência de determinada doença em relação a variáveis referentes à população, como idade ou grupo etário, gênero, ocupação, etnia, entre outras.

A mortalidade se refere ao número de óbitos que ocorreram em determinada população, em determinado período de tempo. Os indicadores de mortalidade, por sua vez, podem ser calculados como a razão entre frequências absolutas de óbitos e números de sujeitos expostos ao risco de morrer.

A letalidade é expressa em porcentagem, é a taxa de proporção entre o número de mortes por uma doença e o número total de doentes acometidos pela doença por determinado período de tempo.

Conhecer e determinar os indicadores de saúde possibilita avaliar e caracterizar o cenário da saúde e, através deles, realizar procedimentos e tomadas de decisões de forma mais consciente e assertiva.



UNIDADE V

INTRODUÇÃO AOS TESTES DE HIPÓTESES



OBJETIVOS

- ✓ Conhecer os conceitos e aplicações dos Testes de Hipóteses;
- ✓ Identificar as hipóteses de teste, através da interpretação de resultados e conclusões baseadas na probabilidade de significância.



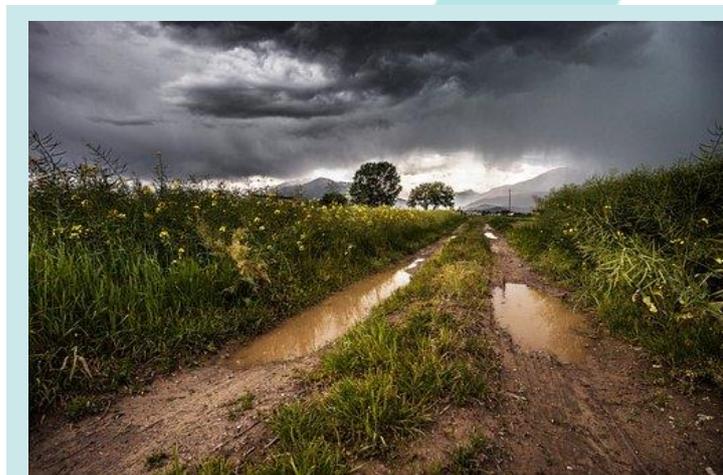
META DA UNIDADE

Um dos principais assuntos da Estatística moderna é a inferência estatística. A inferência estatística é dividida em dois grandes tópicos: a estimação de parâmetros e os testes de hipóteses. No desenvolvimento dos métodos da estatística moderna, as primeiras técnicas de inferência que apareceram foram as que faziam diversas hipóteses sobre a natureza da população da qual se extraíram os dados.

Aqui estudaremos o aspecto da inferência estatística chamado de teste de hipóteses, cujo o objetivo é decidir se uma afirmação, em geral, sobre **parâmetros** de uma ou mais populações é, ou não, apoiado pela evidência obtida de dados amostrais. Tal afirmação é o que se chama **Hipótese Estatística** e a regra usada para decidir se ela é verdadeira ou não é o **Teste de Hipóteses**.

Observe a FIG. 1 abaixo e responda à seguinte pergunta: Será que vai chover?

FIGURA 24 – IMAGEM DE UM DIA NUBLADO



FONTE: [HTTPS://CDN.PIXABAY.COM/PHOTO/2014/07/31/23/47/COUNTRY-ROAD-407207__340.JPG](https://cdn.pixabay.com/photo/2014/07/31/23/47/country-road-407207__340.jpg). ACESSO EM 23/01/2023

Segundo o dicionário de português, a palavra hipótese se refere a uma suposição que se faz acerca de uma coisa possível ou não, da qual se tira uma consequência; teoria provável, admissível, embora ainda não demonstrado.

Em relação à ilustração acima, concluímos nossa teoria baseando-nos no fato de as nuvens estarem bem carregadas.

Mas será que podemos comprovar essa hipótese? Será que essa hipótese pode ser rejeitada? Que elementos temos para aceitar ou rejeitar essa hipótese? Nesta unidade estudaremos como comprovar cientificamente nossas hipóteses.

CONCEITOS BÁSICOS DO TESTE DE HIPÓTESE

Como o teste de hipótese é o foco principal deste capítulo, comecemos com duas definições muito básicas.

DEFINIÇÕES

“Em estatística, uma hipótese é uma reivindicação ou afirmativa sobre uma propriedade de uma população. Um teste de hipótese (ou teste de significância) é um procedimento para o teste de uma afirmativa sobre uma propriedade de uma população. Definição de Hipótese Nula e Hipótese Alternativa”. (Triola, 2017)

A “propriedade” referida na definição precedente é, em geral, o valor de um parâmetro populacional, e mostramos aqui alguns exemplos de hipóteses típicas (ou afirmativas):

- $\mu < 98,6$ °F A temperatura corporal média é menor do que 98,6 °F.
- $p > 0,5$ O método XSORT (programação) de seleção de gênero aumenta a probabilidade de um bebê ser menina, de modo que a probabilidade de uma menina é maior do que 0,5.
- $\sigma = 15$ A população de estudantes de faculdade tem escores de QI com desvio-padrão igual a 15.

As hipóteses estatísticas sempre comparam dois ou mais parâmetros, afirmando que são iguais ou não. Essas hipóteses podem ser:

Hipótese Nula ou de Nulidade (H₀), que estabelece a ausência de diferenças entre os parâmetros. É sempre a primeira a ser formulada.

Hipótese Alternativa (H_a), é a hipótese contrária à hipótese nula. Estabelece a presença de diferenças entre os parâmetros. Geralmente, é a que pesquisador quer ver confirmada.

Para criar hipótese, use os seguintes passos:

Objetivo	
Identificar a hipótese nula e a hipótese alternativa de modo que o teste formal de hipótese inclua esses componentes	padrão que são usados frequentemente em diferentes disciplinas.
Hipótese nula (representada por H₀)	
Afirmativa de que o valor de um parâmetro populacional (como proporção, média ou desvio-padrão) é igual a algum valor especificado (o termo nula é usado para indicar nenhuma mudança, ou nenhum efeito, ou nenhuma diferença.) Testamos a hipótese nula diretamente, no sentido de que assumimos ou supomos que seja verdadeira	chegamos a uma conclusão de rejeitá-la ou deixar de rejeitá-la. Exemplo: Eis uma hipótese nula típica envolvendo uma proporção: H ₀ : $p = 0,5$.
Hipótese alternativa (representada por H₁ ou H_a ou H_A)	
Afirmativa de que o parâmetro tem um valor que, de alguma forma, difere da hipótese nula. Para os métodos deste capítulo, a forma simbólica da hipótese alternativa deve usar um destes símbolos: $<$, $>$, \neq . Exemplo: Eis diferentes exemplos de hipóteses alternativas que envolvem proporções:	$H_1: p > 0,5$ $H_1: p < 0,5$ $H_1: p \neq 0,5$ A afirmativa original pode se tornar a hipótese nula (como na afirmativa de que $p = 0,5$), pode se tornar a hipótese alternativa (como na afirmativa de que $p > 0,5$) ou pode não corresponder nem à hipótese nula nem à hipótese alternativa (como na afirmativa de que $p \geq 0,5$).

TIPOS DE TESTE DE HIPÓTESE: BILATERAL, UNILATERAL À ESQUERDA, UNILATERAL À DIREITA

Em geral, somente a estatística de teste não nos fornece informação suficiente para tomarmos uma decisão sobre a afirmativa que está sendo testada. Para tal decisão, essa abordagem exige que determinemos primeiro se nosso teste de hipótese é bilateral ou unilateral.

A região crítica (ou região de rejeição) corresponde aos valores da estatística de teste que nos levam a rejeitar a hipótese nula. Dependendo da afirmativa em teste, a região crítica poderia estar nas duas caudas extremas, poderia estar na cauda esquerda ou poderia estar na cauda direita.

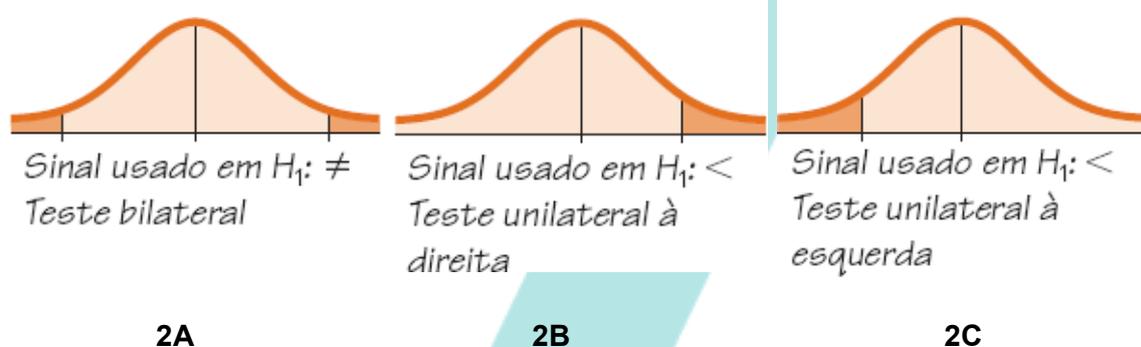
Teste bilateral: a região crítica está nas duas regiões extremas (caudas) sob a curva (como no gráfico superior na FIG. 2a).

Teste unilateral à esquerda: a região crítica está na região extrema (cauda) esquerda sob a curva (como no gráfico do meio na FIG. 2b).

Teste unilateral à direita: a região crítica está na região extrema (cauda) direita sob a curva (como no gráfico inferior na FIG. 2c).

figura 2 - Região Crítica nos testes bilateral, unilateral esquerda e unilateral direita

FIGURA 25 - REGIÃO CRÍTICA NOS TESTES BILATERAL, UNILATERAL ESQUERDA E UNILATERAL DIREITA



FONTE: TRIOLA, 2017, PÁG.382.



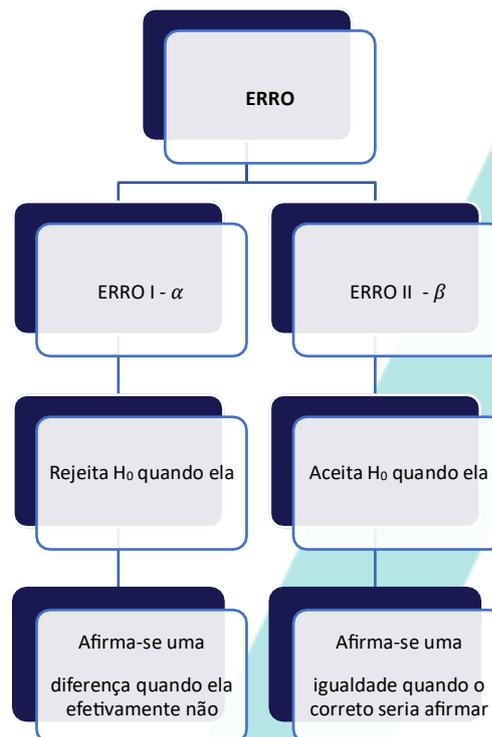
INDICAÇÃO DE VÍDEO

https://youtu.be/l1_RAOehQ9k

TIPOS DE ERROS

Vimos que o processo envolve tomada de decisão em aceitar e rejeitar hipóteses. Sendo assim, há implicações nessa escolha e uma delas é o tipo de erro que se comete ao escolher uma e rejeitar a outra. Vejamos quais são esses tipos de erros e como eles estão associados às hipóteses nula e alternativa.

Figura 26 - Tipos de erros que podem ser cometidos ao se testar hipóteses



Fonte: Elaboração própria, 2023.

No quadro a seguir você pode verificar os erros cometidos de acordo com a decisão tomada pelo pesquisador de aceitar ou não uma hipótese.

TABELA 6 - TIPOS DE ERROS COMETIDOS AO ACEITAR OU REJEITAR UMA HIPÓTESE DE NULIDADE OU ALTERNATIVA

Se a Hipótese Nula (H_0) é			
		VERDADEIRA	FALSA
O PESQUISADOR	ACEITA H_0	DECISÃO CORRETA	COMETE O ERRO TIPO II (β)
	REJEITA H_0	COMETE O ERRO TIPO I (α)	DECISÃO CORRETA

FONTE: ELABORAÇÃO PRÓPRIA, 2023.



INDICAÇÃO DE VÍDEO

<https://youtu.be/cpH2MSRuVp8>

NÍVEL DE SIGNIFICÂNCIA

O nível de significância, também denotado como alfa ou α , é a probabilidade de rejeição da hipótese nula quando ela é verdadeira. Por exemplo, um nível de significância de 0,05 indica um risco de 5% de concluir que existe uma diferença quando não há diferença real.

Nível de significância (α) → probabilidade do cometer Erro I.

NÍVEIS CLÁSSICOS DE SIGNIFICÂNCIA

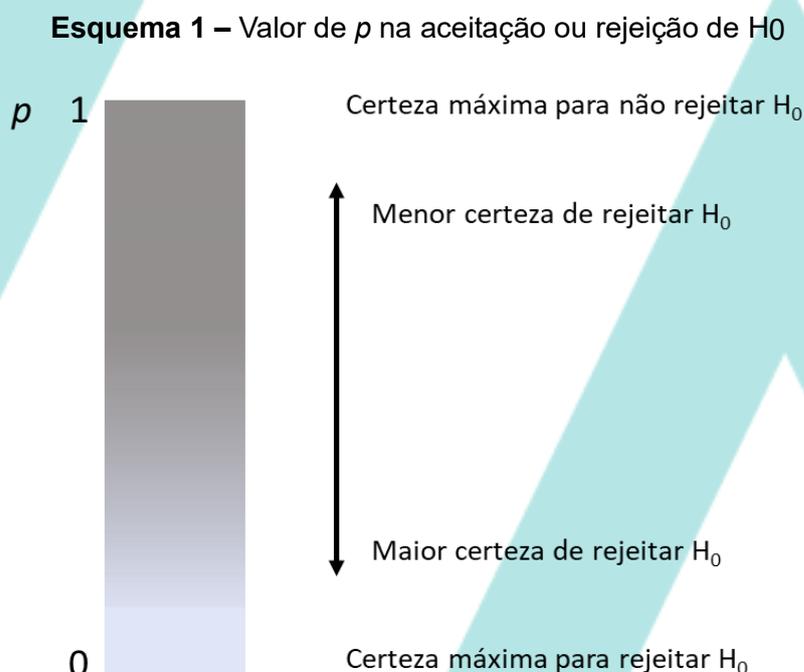
De modo geral, as hipóteses são testadas em três níveis de significância: 1; 5 e 10%. Escolhe-se para apresentar a conclusão o menor valor de α capaz de rejeitar H_0 . Se nenhum dos níveis conseguir rejeitar H_0 , diz-se que o teste é não significativo. Se o menor valor

de alfa que rejeita H_0 é, por exemplo, 5%, diz-se que o teste é significativo ao nível de 5%.

A probabilidade de rejeição de H_0 associada ao resultado da experiência é denominada **nível de significância** do teste, ou simplesmente, p .

- ✓ Quanto menor p , maior a evidência de que existam diferenças, então rejeita-se H_0 com maior certeza;
- ✓ Quanto maior p , maior a evidência de que não existam diferenças, então diminui a certeza da rejeição de H_0 .

As decisões associadas a esses níveis de significância costumam ser classificadas de acordo com o seguinte esquema:



Fonte: Elaboração própria, 2023.

Assim, podemos perceber que, na construção de um teste de hipóteses, procuramos controlar o erro tipo I fixando o valor de α . Entendemos, então, que o nível de significância não é um valor que calculamos a partir da amostra, mas é escolhido quando nosso estudo está sendo planejado. Nas publicações científicas da área da saúde, na grande maioria das vezes, α é fixado em 0,05.

DEFINIÇÃO E INTERPRETAÇÃO DE P-VALOR

Observamos que o resultado de um teste estatístico de hipóteses é a decisão de rejeitar ou não a hipótese nula. A menor possível escolha de α para a qual rejeitamos H_0 a partir da presente amostra equivale ao nível descritivo do teste ou valor p (muitas vezes abreviada somente pela letra p) (TRIOLA, 2017).

Por exemplo, se temos $p = 0,048$, isso significa que rejeitamos H_0 para escolhas de α maiores ou igual a 0,048 (inclusive para $\alpha = 0,05$).

Isso nos leva a uma regra prática segundo a qual rejeitamos H_0 menor que o nível de significância escolhido, e não rejeitamos H_0 se o valor p é caso contrário.

Observe que, em nenhum instante, estamos modificando o nível de significância previamente estabelecido. Se previamente estabelecemos α em, por exemplo, “0,05”, esse valor não se modifica na análise dos dados. A obtenção do valor p consiste somente em entender qual seria o menor valor que precisaríamos ter atribuído a α para que a hipótese nula fosse rejeitada a partir de nossa amostra. Se em um teste de hipóteses encontramos, por exemplo, um valor p igual a 0,78, entendemos que, para que a hipótese nula fosse rejeitada, precisaríamos ter estabelecido previamente um nível de significância maior ou igual a 0,78. Porém, se o nível de significância foi estabelecido em 0,05 (ou seja, menor que 0,78), entende-mos que a hipótese nula não é rejeitada (TRIOLA, 2017).

É importante observar que o resultado de um teste de hipóteses não é o valor p em si, mas a decisão de rejeitar ou não uma hipótese nula de interesse. Em adição, observamos que, enquanto o nível de significância é determinado pelo pesquisador, o valor p é obtido dos resultados de uma amostra. Portanto, o valor p é uma quantidade que se modifica de uma amostra para outra (TRIOLA, 2017).



INDICAÇÃO DE VÍDEO

<https://youtu.be/8t9PID7S5zk>

ERRO TIPO 2 E PODER DO TESTE

Já sabemos que a probabilidade de se rejeitar H_0 quando ela é verdadeira foi definido como nível de significância α . Entretanto, a probabilidade de se rejeitar H_1 quando ela é verdadeira é a probabilidade de cometer o **Erro II**, representado pela letra grega Q e seu complemento $(1 - \beta)$, o que se conhece como poder (ou potência) de um teste. Assim,

poder de um teste é a probabilidade de rejeitar H_0 quando ela é falsa.

Contudo, notamos que a probabilidade Q não é um número fixo, visto que existem infinitas possibilidades para μ_0 (medida ou valor de referência) quando H_A é verdadeira. Assim, Q é uma função de μ , melhor reescrita na forma $b(\mu)$

$$1 - \beta(\mu) = P(\text{rejeitar } H_0 | n),$$

ou seja, a probabilidade de rejeitar a hipótese nula para um valor particular de μ é chamada **função poder** do teste de hipóteses.

A função poder é muito utilizada na pesquisa em saúde para a determinação do tamanho amostral necessário ao estudo. Nesse caso, primeiro fixamos α em um valor (geralmente 0,05). Em seguida, fixamos a função poder $1 - \beta(\mu)$ para um valor particular de μ_0 . Valores comuns para $1 - \beta(\mu)$ são 0,05, 0,1 ou 0,2 (TRIOLA, 2017).

Nas publicações em saúde, a função poder não costuma ultrapassar 0,2.

O poder é uma característica do teste de hipóteses, e não da amostra ou de nossos dados. Portanto, não é correto falarmos em “poder da amostra”.

RESUMINDO

O teste de hipótese é um procedimento de inferência estatística usual quando desejamos, por meio de uma amostra, estudar e inferir sobre uma população. Entretanto, os testes de hipóteses têm algumas características importantes quanto à interpretação de seus resultados.

Em um teste estatístico de hipóteses, consideramos duas hipóteses:

- Hipótese alternativa: denotada por H_A aquilo que ele deseja verificar.
- Hipótese nula: denotada por H_0 , é a “hipótese do pesquisador”, é o complemento da hipótese alternativa.

A hipótese alternativa leva esse nome uma vez que, em muitas aplicações, ela expressa uma alternativa ao que já se acredita a respeito do parâmetro.

Com base nos resultados de uma amostra aleatória de tamanho n , tomamos a decisão de rejeitar ou não rejeitar a hipótese nula H_0 . O teste de hipóteses, em princípio, considera a hipótese nula como verdadeira. Estaremos, assim, sujeitos a dois tipos de erros:

- Erro tipo I: se rejeitamos H_0 , mas H_0 é verdadeira, cometemos um erro tipo I.
- Erro tipo II: se não rejeitamos H_0 , mas H_0 é falsa, cometemos um erro tipo II.

A probabilidade de se cometer um erro tipo I é chamada nível de significância do teste e é denotada pela letra grega α (alfa) e o poder de teste é quando rejeitamos

H_1 quando ela é verdadeira, ou seja, é a probabilidade de cometer o **Erro II**.



UNIDADE VI

MEDIDAS DE EFEITO



OBJETIVOS

- ✓ Analisar dados de medidas de efeitos, cálculos, aplicações e interpretações.



METAS DA UNIDADE

Geralmente, pesquisas que envolvem estudos epidemiológicos têm como objetivo identificar associação e efeito entre fatores de risco ou de proteção e doenças ou medidas clínicas. A relação entre um fator de exposição e um desfecho pode ser expressa por meio de medidas que quantificam a dimensão dessa associação. O tipo de medida de associação ou efeito mais adequada para avaliar esta relação é definida pelo delineamento do estudo e o tipo das variáveis estudadas.

Por exemplo, se o delineamento é coorte ou experimental, a medida de associação indicada é o RR (RR), chamado também de razão de riscos. O RR é a razão das medidas de incidência de doenças em duas populações. Incidência é a proporção dos indivíduos da população em risco no início de um intervalo de tempo que se tornaram novos casos da doença até o fim do intervalo. No caso de um delineamento transversal, a associação entre as variáveis é dada pela razão de prevalência (RP), em que a medida de frequência estimada é a prevalência pontual. Prevalência é a fração de um grupo de pessoas que possui uma condição ou desfecho clínico em um dado ponto do tempo. A interpretação da RP é similar ao do RR, considerando a prevalência ao invés da incidência.

Nesta unidade, daremos ênfase ao estudo das medidas de associação e de efeito: RR, Razão das Chances e Razão das Chances de Prevalência, destacando suas definições e principais características.

RAZÃO DE RISCO (RR)

Alguma vez você já se perguntou: “Qual probabilidade de as pessoas expostas desenvolverem a doença é quantas vezes maior em relação às pessoas não expostas?”. Para responder a essa questão, usa-se o RR ou razão de riscos, ou seja, a razão entre a incidência em pessoas expostas e a incidência em pessoas não expostas, colocando em termos

matemáticos, temos:

$$RR = \frac{\text{incidência em pessoas expostas}}{\text{incidência em pessoas não expostas}}$$

Perceba que o numerador e o denominador da expressão matemática são duas taxas. Dessa forma, podemos dizer que o RR é a razão entre duas taxas de incidência.

Um RR igual a 1 indica que a incidência do desfecho foi igual nos dois grupos comparados, ou seja, a exposição a determinado fator não causou maior risco de doença nos expostos comparativamente em relação aos não expostos. Um RR de 1,5 significa que o risco entre os expostos foi 50% maior do que entre os não expostos $[(RR - 1) \times 100\%]$. Por outro lado, um RR menor do que 1 indica que a exposição foi um fator de proteção. Por exemplo, um RR de 0,7 demonstra que houve uma proteção entre os expostos de 30% $[(1 - RR) \times 100\%]$. Resumidamente temos na Tabela 1 as possibilidades de interpretação do RR de uma doença.

TABELA 7 - INTERPRETAÇÃO DO RISCO RELATIVO (RR) DE UMA DOENÇA

Se RR = 1	Risco nos expostos igual ao risco nos não expostos (sem associação)
Se RR > 1	Risco nos expostos maior que o risco nos não expostos (associação positiva; possivelmente causal)
Se RR < 1	Risco nos expostos menor que o risco nos não expostos (associação negativa; possivelmente protetora)

FONTE: GORDIS, 2017, P.217.

RR NOS PRINCIPAIS TIPOS DE ESTUDO

O RR é calculado, diretamente, nos estudos que geram taxas de incidência: é o caso dos ensaios clínicos e dos estudos de coorte.

Em estudos de casos-controles, o RR não pode ser diretamente computado; ele é estimado, indiretamente, pelo cálculo do que os autores ingleses denominam de *Odds Ratio* (OR) ou Razão de Chances.

Por aproximação, em estudos transversais, uma estimativa do RR é obtida a partir de coeficientes em expostos e não-expostos; ou calcula-se o OR ou a razão de prevalências.

PROBABILIDADE E CHANCE

Probabilidade e chance são termos usados como sinônimos no dia-a-dia. Os livros de estatística fazem uma sutil diferença entre eles: a probabilidade compara o número de casos favoráveis com o de casos possíveis, enquanto a chance, o número de casos favoráveis com o de casos desfavoráveis. Trata-se, portanto, de diferentes maneiras de exprimir as possibilidades de ocorrência de um mesmo evento.

Exemplo: probabilidade x chance

A probabilidade de sair cara quando se joga uma moeda é $1/2$ (ou 0,5). A chance de sair cara é de 1 para 1.

A probabilidade de jogar um dado e sair o número 6 é $1/6$. A chance de sair o número 6 é de 1 para 5.

Cálculo do RR em Estudos de Coorte

Em um estudo de coorte, o RR pode ser calculado diretamente. Vamos usar um exemplo hipotético que envolvem pesquisa sobre tabagismo e doença coronariana (CHD).

A Tabela 2 mostra um estudo de coorte hipotético de 3.000 fumantes e 5.000 não fumantes, para investigar a relação entre fumo e desenvolvimento de doença coronariana (CHD) por um período de 1 ano.

TABELA 8 - FUMO E DOENÇA CORONARIANA (CHD): ESTUDO DE COORTE HIPOTÉTICO

	Desenvolvera m CHD	Não Desenvolvera m CHD	Totais	Incidência por 1.000 pessoas/ano
Fumantes	84	2.916	3.000	28,0
Não Fumantes	87	4.913	5.000	17,4

FONTE: GORDIS, 2017, P.218.

O cálculo do RR para o delineamento do estudo de coorte é apresentado na Tabela 3.

TABELA 9 - CÁLCULO DE RISCO EM ESTUDO DE COORTE

Acompanhamento para ver se				
	A doença se desenvolveu	A doença não se desenvolveu	Totais	Taxa de incidência da doença
<i>Primeira Seleção</i> { <i>Expostos</i> <i>Não expostos</i>	a	b	a + b	$\frac{a}{a + b}$
	c	d	c + d	$\frac{c}{c + d}$
	$\frac{a}{a + b} = \text{Incidência nos expostos}$			$\frac{c}{c + d} = \text{Incidência nos não expostos}$

FONTE: GORDIS, 2017, P.218.

A partir das informações da tabela, calculamos o RR

$$RR = \frac{\text{incidência em pessoas expostas}}{\text{incidência em pessoas não expostas}} = \frac{\left(\frac{a}{a + b}\right)}{\left(\frac{c}{c + d}\right)}$$

Aplicando os valores do exemplo, temos:

$$RR = \frac{\left(\frac{84}{84 + 2916}\right)}{\left(\frac{87}{87 + 4913}\right)}$$

$$RR = \frac{\left(\frac{84}{3000}\right)}{\left(\frac{87}{5000}\right)} = \frac{28}{17,4} = 1,61$$

Isso mostra que os expostos, ou seja, os fumantes, têm 61% de probabilidade a mais do que os não fumantes (não expostos) de desenvolver doença coronariana (CDH).

RAZÃO DAS CHANCES OU *ODDS RATIO*

Acabamos de ver que, para calcular o RR, é necessário ter valores para a incidência da doença em expostos e não expostos, que podem ser obtidas em um estudo de coorte.

Em um estudo de casos-controle, no entanto, não conhecemos a incidência na população exposta e não exposta, pois começamos com pessoas doentes (casos) e não doentes (controles). Por essa razão, em estudos de casos-controle, não podemos calcular o RR diretamente.

Nesse caso, outra medida de associação, o **Odds Ratio (OR)**, pode ser obtido em estudos de coorte ou de casos-controle e pode ser usado no lugar do RR.

O **OR** pode ser calculado a partir da seguinte fórmula:

$$OR = \frac{ad}{bc}$$

Ou seja, é resultado do produto cruzado entre os valores obtidos na pesquisa.

Historicamente, o **OR** foi utilizado como aproximação do RR. No entanto, mais recentemente, constatou-se que o **OR** constitui também uma medida própria de risco, inclusive com vantagens sobre o RR. Uma delas é que, ao contrário do RR, o **OR** pode ser calculado em qualquer tipo de estudo epidemiológico. Daí o seu uso estar se expandindo, embora o tema esteja ainda em franco debate e desenvolvimento.

A Tabela 10 apresenta uma relação entre a fórmula do RR e a do **OR**, ainda considerando o exemplo hipotético apresentado na subseção 6.1.1.

TABELA 10 - CÁLCULO DE RISCO EM ESTUDO DE COORTE, COM DADOS HIPOTÉTICOS

Risco Relativo	$RR = \frac{\text{incidência em pessoas expostas}}{\text{incidência em pessoas não expostas}} = \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)}$ $RR = \frac{\left(\frac{84}{3000}\right)}{\left(\frac{87}{5000}\right)} = \frac{28}{17,4} = 1,61$
Odds Ratio	$OR = \frac{ad}{bc} = \frac{84 \cdot 4913}{2916 \cdot 87} = \frac{412692}{253692} = 1,63$

FONTE: ELABORAÇÃO PRÓPRIA, 2023.

O RR e o OR sempre apontam para a mesma direção, embora possa haver diferenças numéricas entre as duas medidas. Tais diferenças são mínimas, quando as frequências das doenças são baixas, e se tornam mais evidentes em doenças de maior frequência.

- **Interpretação do OR:** Interpretamos o OR da mesma forma como interpretamos o RR. Veja a Tabela 5.

TABELA 11 - INTERPRETAÇÃO DOS RESULTADOS DO OR

Se OR = 1	Indica que a exposição não está associada à doença.
Se OR > 1	Indica que a exposição está associada com a doença positivamente.
Se OR < 1	Indica que a exposição está associada com a doença negativamente.

FONTE 12: ELABORAÇÃO PRÓPRIA, 2023.

- **Quando o OR é uma boa estimativa do RR?**

Em um estudo de casos-controle, só o **OR** pode ser calculado como medida de associação, enquanto em um estudo de coorte, tanto **RR** quanto **OR** são válidos como medidas de associação. Contudo, muitas pessoas sentem-se mais confortáveis usando o **RR**, que é a medida de associação mais frequentemente relatada na literatura quando resultados de estudos de coorte são publicados.

Mesmo quando o **OR** é usado, as pessoas frequentemente estão interessadas em saber quanto se aproxima do **RR**.⁹

Quando o OR, obtido em estudos de casos-controle, é uma boa aproximação do RR na população? Quando as três condições seguintes são satisfeitas:

1. Quando os casos estudados forem representativos, em relação à história de exposição, de todas as pessoas com a doença na população de onde saíram os casos.
2. Quando os controles estudados forem representativos, em relação à história de exposição, de todas as pessoas sem a doença na população de onde saíram os casos.
3. Quando a doença estudada não ocorre frequentemente.

A comparação apresentada na Tabela 4 acima mostra uma situação hipotética em que o **OR** é uma boa estimativa de RR quando a doença não é frequente. Agora, se os valores são discrepantes, o OR não é uma boa estimativa de RR quando a doença é frequente. Veja a FIG. 26.

FIGURA 26 - EXEMPLO: O OR NÃO É UMA BOA ESTIMATIVA DE RR QUANDO A DOENÇA É FREQUENTE

	Desenvol- veram doença	Não desenvol- veram doença	
Expostos	50	50	100
Não expostos	25	75	100

Risco relativo = $\frac{50/100}{25/100} = 2$

Odds ratio = $\frac{50 \times 75}{25 \times 50} = 3$

FONTE: GORDIS, 2017, P. 223.

Vemos, pois, que o OR é, por si só, uma medida de associação válida, mesmo sem considerarmos o RR. Se, entretanto, escolhermos usar o RR como indicador de associação, quando a ocorrência da doença não for frequente, o OR será uma ótima aproximação do RR.



PARA SABER

O odds ratio (razão de chances) é uma medida útil de associação, por si só, tanto nos estudos de coorte como nos de casos-controlle.

Em um estudo de coorte, o risco relativo pode ser calculado diretamente.

Em um estudo de casos-controlle, o risco relativo não pode ser calculado diretamente, assim o odds ratio (razão de produtos cruzados) é usado como estimativa do risco relativo, quando o risco da doença é baixo.

RAZÃO DE PREVALÊNCIAS

Em estudos nos quais se mede prevalência ao invés de incidência (como no caso de estudos transversais), pode-se usar a razão de prevalências (RP) como medida de associação.

Há também a possibilidade de fazer comparações entre duas prevalências categorias diferentes como por exemplo: comparação da prevalência do sexo masculino com do sexo feminino.

A comparação entre duas medidas de prevalências dá origem a uma medida de associação denominada Razão de Prevalência que, por sua vez, não mede o risco, visto que prevalência mede a probabilidade de se ter a doença e não o risco de se desenvolver a doença (que é a incidência).

Dessa forma é correto dizer que RP não indica a magnitude do risco, pois a prevalência é fruto de estudo transversal. Ou seja, o risco não pode ser evidenciado pelo fato de as características da exposição serem coletadas ao mesmo tempo do desfecho.

Assim, o que se tem é que a RP indica magnitude de associação entre duas prevalências e sugere ou gera hipótese de existência de risco.

A RP é obtida por:

$$RP = \frac{\textit{prevalência dos expostos}}{\textit{prevalência dos não expostos}}$$

Interpretação dos resultados da RP: A interpretação dos resultados da razão de prevalência

TABELA 12 - INTERPRETAÇÃO DOS RESULTADOS DA RAZÃO DE PREVALÊNCIA

Se RP = 1	indica que a prevalência do evento de interesse é igual nos dois grupos comparados; portanto, a exposição não está associada ao desfecho.
Se RP > 1 (Associação Positiva)	maior probabilidade de doença entre os expostos. Indica que a exposição está associada ao desfecho (evento de interesse).
Se RP < 1 (Associação Inversa)	menor probabilidade de doença entre os expostos. Indica que a exposição está associada ao desfecho (evento de interesse).

FONTE 13: ELABORAÇÃO PRÓPRIA, 2023.

OUTRA DEFINIÇÃO PARA RAZÕES DE PREVALÊNCIAS

Essa definição é dada considerando as chances de prevalência bruta (PO) que, por sua vez, é o produto da taxa de incidência bruta (I) pela duração média da doença (\bar{D}), quando a população em risco e o grupo de prevalência são estacionários e não há migração para dentro ou para fora do grupo de prevalência.

Usando essa relação, separadamente para uma só população, sob exposição e sob não exposição, ou para uma população exposta e uma não exposta, temos:

$$PO_1 = I_1 \bar{D}_1$$

em que os subscritos 1 e 0 referem-se a expostos e a não expostos, respectivamente. Se a duração média da doença for a mesma, independente da exposição, isto é, se $\bar{D}_1 = \bar{D}_0$, a razão de prevalências bruta será igual à razão da taxa de incidência bruta:

$$RP = \frac{PO_1}{PO_0} = \frac{I_1}{I_0} = I$$

Se a exposição afetar a mortalidade, também alterará a distribuição etária do potencial. Assim, como as pessoas mais velhas tendem a morrer mais cedo, a exposição afetará indiretamente a duração média, de modo que D_1 não será igual a D_0 . Nesse caso, a equação anterior não se manterá exatamente, embora ainda possa se manter aproximadamente.

Vantagens do estudo de prevalência: A escolha por delineamentos de estudos transversais com a identificação da prevalência representam uma economia considerável de recursos em comparação aos estudos de coorte.

Dessa vantagem decorre ser necessário avaliar a prevalência da doença ou outro evento de interesse em um ponto no tempo, ao invés do que procurar continuamente por incidentes casos durante um longo período de tempo.

Em contrapartida, ocorre um enfraquecimento nas inferências, uma vez que pode ser muito mais difícil entender a relação temporal entre várias exposições e a ocorrência de doenças. Assim, é importante sempre considerar os objetivos da pesquisas e fazer a escolha correta das ferramentas de análise.



RESUMO DA UNIDADE

Nesta unidade estudamos as medias de efeitos expressas como razão, sendo elas o Risco (RR), a Razão de Prevalência (RP) e *Odds Ratio* (OR). Todas elas são calculadas através da razão entre a medida de ocorrência de um desfecho entre expostos pela medida de ocorrência do mesmo desfecho nos não- expostos, fornecendo um valor que expressa a força de associação existente entre o fator de exposição e o desfecho em estudo.

Essas medidas podem variar entre 0 e + 8. Quando apresentam valor igual a 1, dizemos que as frequências de doenças entre expostos e não expostos são iguais, pois o numerador e o denominador são idênticos e, conseqüentemente, não existe associação entre exposição e desfecho. Valores maiores que 1 (frequência da doença maior nos expostos em relação aos não- expostos) sugerem que a exposição seja um fator de risco e valores menores de 1 (frequência de doença menor nos expostos em relação aos não expostos) sugerem que a exposição seja um fator de proteção.



REFERÊNCIAS

VIEIRA, Sonia. **Estatística básica** – 2ª edição revista e ampliada. São Paulo: Cengage Learning Brasil, 2018. E-book. ISBN 9788522128082. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788522128082/> . Acesso em: 11 jan. 2023.

TRIOLA, Mario F. **Introdução à Estatística**, 12ª edição. Rio de Janeiro: Grupo GEN, 2017. E-book. ISBN 9788521634256. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521634256/> . Acesso em: 11 jan. 2023.

MARTINS, Amanda de Ávila B.; TEIXEIRA, Deborah; BATISTA, Bruna G.; STEFFENS, Daniela. **Epidemiologia**. Porto Alegre: Grupo A, 2018. E-book. ISBN 9788595023154. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595023154/> . Acesso em: 23 jan. 2023.

ROUQUAYROL, Maria Z.; GURGEL, Marcelo. Rouquayrol - **Epidemiologia e saúde**. Rio de Janeiro : MedBook Editora, 2017. E-book. ISBN 9786557830000. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786557830000/> . Acesso em: 23 jan. 2023.

PEREIRA, Maurício G. **Epidemiologia - Teoria e Prática**. São Paulo: Grupo GEN, 1995. E-book. ISBN 9788527736077. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788527736077/> . Acesso em: 11 jan. 2023.

GOMES, E. C. de S.. **Conceitos e ferramentas da epidemiologia** . Recife: Ed. Universitária da UFPE, 2015. 83 p. Disponível em: https://ares.unasus.gov.br/acervo/bitstream/ARES/3355/1/3con_ferra_epidemiolo_2016-2.pdf . Acesso em: 31/01/2023.

VIEIRA, Sonia. **Estatística básica** – 2ª edição revista e ampliada. São Paulo: Cengage Learning Brasil, 2018. E-book. ISBN 9788522128082. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788522128082/> . Acesso em: 11 jan. 2023.

TRIOLA, Mario F. **Introdução à Estatística**, 12ª edição. Rio de Janeiro: Grupo GEN, 2017. E-book. ISBN 9788521634256. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521634256/> . Acesso em: 31 jan. 2023.

MARTINS, Amanda de Ávila B.; TEIXEIRA, Deborah; BATISTA, Bruna G.; STEFFENS, Daniela. **Epidemiologia**. Porto Alegre: Grupo A, 2018. E-book. ISBN 9788595023154. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595023154/> . Acesso em: 23 jan. 2023.

ROUQUAYROL, Maria Z.; GURGEL, Marcelo. Rouquayrol - **Epidemiologia e saúde**. Rio de Janeiro : MedBook Editora, 2017. E-book. ISBN 9786557830000. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786557830000/> . Acesso em: 23 jan. 2023.

PEREIRA, Maurício G. **Epidemiologia - Teoria e Prática**. São Paulo: Grupo GEN, 1995. E-book. ISBN 9788527736077. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788527736077/> . Acesso em: 31 jan. 2023.

VIEIRA, Sonia. Estatística básica – 2ª edição revista e ampliada. São Paulo: Cengage Learning Brasil, 2018. E-book. ISBN 9788522128082. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788522128082/> . Acesso em: 15 jan. 2023.

TRIOLA, Mario F. Introdução à Estatística, 12ª edição. Rio de Janeiro: Grupo GEN, 2017. E-book. ISBN 9788521634256. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521634256/> . Acesso em: 15 jan. 2023.

VIEIRA, Sonia. Fundamentos de Estatística, 6ª edição. São Paulo: Grupo GEN, 2018. E-book. ISBN 9788597019315. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788597019315/> . Acesso em: 15 jan. 2023.

ARANGO, Hector G. Bioestatística - Teórica e Computacional, 3ª edição. Rio de Janeiro: Grupo GEN, 2009. E-book. ISBN 978-85-277-1943-8. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/978-85-277-1943-8/> . Acesso em: 15 jan. 2023.

PEREIRA, Maurício G. Epidemiologia - Teoria e Prática. São Paulo: Grupo GEN, 1995. E-book. ISBN 9788527736077. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788527736077/> . Acesso em: 15 jan. 2023.

MARTINS, Amanda de Ávila B.; TEIXEIRA, Deborah; BATISTA, Bruna G.; STEFFENS, Daniela. **Epidemiologia**. Porto Alegre: Grupo A, 2018. E-book. ISBN 9788595023154. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595023154/> . Acesso em: 23 jan. 2023.

PEREIRA, Maurício G. **Epidemiologia - Teoria e Prática**. São Paulo: Grupo GEN, 1995. E-book. ISBN 9788527736077. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788527736077/> . Acesso em: 15 jan. 2023.

FILHO, Naomar de A.; BARRETO, Mauricio L. **Epidemiologia & Saúde - Fundamentos, Métodos e Aplicações**. Rio de Janeiro: Grupo GEN, 2011. E-book. ISBN 978-85-277-2119-6. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/978-85-277-2119-6/> . Acesso em: 23 jan. 2023.

ROUQUAYROL, Maria Z.; GURGEL, Marcelo. Rouquayrol - **Epidemiologia e saúde**. Rio de Janeiro : MedBook Editora, 2017. E-book. ISBN 9786557830000. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786557830000/> . Acesso em: 23 jan. 2023.

GORDIS, Leon. **Epidemiologia**. Rio de Janeiro: Thieme Brazil, 2017. E-book. ISBN 9788567661926. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788567661926/> . Acesso em: 23 jan. 2023.

VIEIRA, Sonia. Estatística básica – 2ª edição revista e ampliada. São Paulo: Cengage Learning Brasil, 2018. E-book. ISBN 9788522128082. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788522128082/> . Acesso em: 15 jan. 2023.

TRIOLA, Mario F. Introdução à Estatística, 12ª edição. Rio de Janeiro: Grupo GEN, 2017. E-book. ISBN 9788521634256. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788521634256/> . Acesso em: 15 jan. 2023.

VIEIRA, Sonia. Fundamentos de Estatística, 6ª edição. São Paulo: Grupo GEN, 2018. E-book. ISBN 9788597019315. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788597019315/> . Acesso em: 15 jan. 2023.

ARANGO, Hector G. Bioestatística - Teórica e Computacional, 3ª edição. Rio de Janeiro: Grupo GEN, 2009. E-book. ISBN 978-85-277-1943-8. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/978-85-277-1943-8/> . Acesso em: 15 jan. 2023.

PEREIRA, Maurício G. Epidemiologia - Teoria e Prática. São Paulo: Grupo GEN, 1995. E-book. ISBN 9788527736077. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788527736077/> . Acesso em: 15 jan. 2023.

ROTHMAN, Kenneth; GREENLAND, Sander; LASH, Timothy. Epidemiologia moderna. Porto Alegre: Grupo A, 2011. E-book. ISBN 9788536325880. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788536325880/> . Acesso em: 30 jan. 2023.

GORDIS, Leon. **Epidemiologia**. Rio de Janeiro: Thieme Brazil, 2017. E-book. ISBN 9788567661926. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788567661926/> . Acesso em: 23 jan. 2023.

MARTINS, Amanda de Ávila B.; TEIXEIRA, Deborah; BATISTA, Bruna G.; STEFFENS, Daniela. **Epidemiologia**. Porto Alegre: Grupo A, 2018. E-book. ISBN 9788595023154. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595023154/> . Acesso em: 23 jan. 2023.

FLETCHER, Grant S. Epidemiologia Clínica: Elementos Essenciais. Porto Alegre : Grupo A, 2021. E-book. ISBN 9786558820161. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786558820161/> . Acesso em: 30 jan. 2023.

ROUQUAYROL, Maria Z.; GURGEL, Marcelo. Rouquayrol - **Epidemiologia e saúde**. Rio de Janeiro : MedBook Editora, 2017. E-book. ISBN 9786557830000. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786557830000/> . Acesso em: 23 jan. 2023.