

CONCEITOS INICIAIS

ESTATÍSTICA: ciência que tem objetivo de orientar a coleta, o resumo, a apresentação, a análise e a interpretação de dados.

BIOESTATÍSTICA: aplicação dos métodos estatísticos à solução de problemas biológicos e médicos.

EPIDEMIOLOGIA: ciência que estuda quantitativamente a distribuição dos fenômenos de saúde/doença, e seus fatores condicionantes e determinantes, nas populações humanas.

ORTOPEDIA E TRAUMATOLOGIA: especialidade médica que cuida do trauma do aparelho músculo- esquelético e esta relacionado com o aparelho locomotor.

SAÚDE BASEADA EM EVIDÊNCIAS: é uma abordagem que utiliza as ferramentas da Epidemiologia Clínica; da Estatística; da Metodologia Científica; e da Informática para trabalhar a pesquisa; o conhecimento; e a atuação em Saúde, com o objetivo de oferecer a melhor informação disponível para a tomada de decisão nesse campo.

FATOR: denominada variável preditora ou de exposição, pode ser considerada a possível “causa” de algum efeito.

DESFECHO: resultado final de interesse de um estudo, também denominado de efeito.

POPULAÇÃO: é um termo estatístico amplo utilizado como um conjunto de unidades experimentais ou observacionais que apresenta uma ou mais características em comum e depende do objetivo do estudo.

AMOSTRA: é qualquer fração da população que represente o que desejamos analisar para a realização da pesquisa. É uma parte do todo.

VARIÁVEL: é toda característica que pode ser observada em uma unidade experimental e varia de indivíduo para outro, ex.: raça, sexo, profissão, idade...

DADOS: são informações numéricas ou não e que são obtidos de uma unidade experimental ou observacional resultantes de uma variável.

PARÂMETRO: é o resumo de dados da população.

ESTIMATIVA: é o resumo dos dados da amostra.

CONCEITOS APLICADOS

1) DIVISÃO DA ESTATÍSTICA:

- **DESCRITIVA:** é o resumo e a apresentação dos dados (a parte).

- **INFERENCIAL:** é um conjunto maior de dados (populações) quando apenas partes desses conjuntos (as amostras) foram estudadas (o todo).

2) TIPOS DE VARIÁVEIS:

- **QUANTITATIVAS:** são aquelas que os dados são numéricos e classificam em:

DISCRETAS: são aquelas em que os dados somente podem apresentar determinados valores, números inteiros e por meio de contagem.

CONTÍNUAS: são aquelas cujos dados podem apresentar qualquer valor dentro de um intervalo de variação, são números medidos.

- **QUALITATIVAS:** são as que fornecem dados de natureza não numérica, como cor, raça, sexo; são variáveis categóricas ou atributos.

NOMINAL: não existe uma relação de intensidade entre as categorias, como masculino e feminino e são divididas em dicotômicas que representam 02 categorias e politômicas acima de duas.

ORDINAL: pode existir relação em diferentes graus de intensidade entre elas podendo apresentar até 10 gradações.

3) MEDIDAS DE OCORRÊNCIA:

- **PREVALÊNCIA:** representa a proporção de indivíduos com uma determinada condição em um determinado momento de tempo. Representa casos antigos e novos.

$$\text{EX: } P = \frac{\text{N}^\circ \text{ de casos conhecidos}}{\text{N}^\circ \text{ da população total}} \times 100\%$$

- **INCIDÊNCIA: (TAXA DE INCIDÊNCIA):** número de novos casos surgidos numa determinada população e num determinado intervalo de tempo. Somente casos novos.

$$\text{EX: } i = \frac{\text{N}^\circ \text{ de casos novos}}{\text{N}^\circ \text{ da população sob risco}} \times 100\%$$

4) VIESES:

- **VIÉS DE SELEÇÃO:** Resulta da escolha de um grupo controle que não é representativo da população de referência (de onde surgem os casos). As distorções nos resultados de um estudo que resultam de falhas nos procedimentos utilizados para a seleção dos participantes e/ou de fatores que influenciam a participação na pesquisa.

- **VIÉS DE AFERIÇÃO:** ocorre quando os instrumentos utilizados não medem o valor real das variáveis. Há distorções nos resultados de um estudo decorrentes de erros de mensuração/aferação dos dados da exposição e/ou desfecho.

- **VIÉS DE CONFUSÃO:** Erro sistemático que ocorre quando os resultados de uma associação entre dois fatores (causa-efeito) podem ser imputados, total ou

parcialmente, a um **terceiro fator** não levado em consideração (variável de confusão ou interveniente).

5) VALIDADE:

- **INTERNA:** é a aproximação das inferências à verdadeira relação causa-efeito observada e só se aplica a estudos cujo objetivo é demonstrar relações causais. Representa a implementação do estudo realizado.

- **EXTERNA:** quando ela permite ao pesquisador generalizar os resultados obtidos à outras populações, outros contextos. E vai depender de se poder mostrar que os resultados obtidos nesta pesquisa não são dependentes da amostra. Representa o delineamento do estudo apresentado.

6) DELINEAMENTO DA PESQUISA:

- **OBSERVACIONAIS:** as pesquisas baseiam-se no registro da ocorrência natural de características ou propriedades dos indivíduos e não sofre influência do pesquisador. Subdividem em:

Transversal: a causa e o efeito (fator – desfecho) são avaliados no mesmo momento do tempo. Isso resulta em medidas de prevalência e verifica se existe relação entre as variáveis e são calculadas juntas.

Caso-Controle: utilizado em doenças que não são freqüentes, inicia sempre pelo o Desfecho. Compara um grupo doente (os casos) com um grupo sem a doença (o controle); as variáveis são calculadas separadamente e comparam os resultados finais. Não se geram medidas de prevalência e nem de incidência.

Coorte: sempre inicia pela causa da doença (fator); observam-se por um período longo um grupo tratado e outro grupo de controle, após comparam as taxas de incidência da doença.

- **EXPERIMENTAIS:** abrange os chamados estudos de intervenção que apresentam como característica principal o fato de o pesquisador controlar as condições do experimento do estudo, visando testar hipóteses etiológicas ou avaliar a efetividade dos procedimentos diagnósticos. Subdividem em:

Ensaio Clínico Randomizado(ECTR): estudo com grupos alocados a partir de um processo aleatório de escolha, buscando-se uma distribuição equilibrada de variáveis de confusão; realizado em Humanos. Com possibilidade de dois delineamentos:

ECTR vs mascaramento: o cegamento ocorre alguém que participa do estudo (paciente ou pesquisador) não sabe o que está sendo administrado (efeito placebo). Isso gera um controle de viés de aferição. Podendo ser: uni-cego (paciente não sabe); duplo-cego (paciente e pesquisador não sabem).

ECTR vs Intention-to-treat(ITT): é uma metodologia em que os pacientes são conservados no grupo que foram designados independente da realização ou não da intervenção. Corresponde também um método de análise para ensaios clínicos randomizados, no qual todos os pacientes alocados para um determinado tratamento são analisados juntos neste grupo, independente se completaram ou não o tratamento

Estudos Experimentais em Animais: objetivo da pesquisa experimental é gerar dados que serão transportados para os seres humanos, sendo largamente utilizada no teste de novas drogas e procedimentos. Muito utilizado na parte Biomédica, porém com restrições Bioéticas.

7) TIPOS DE AMOSTRAGEM:

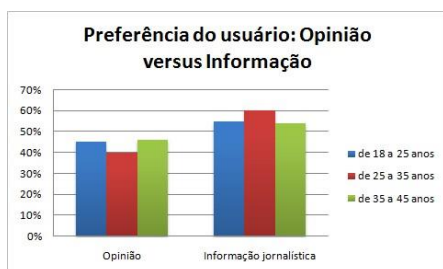
- **AMOSTRAGEM POR CONVENIÊNCIA:** Um grupo da população que é avaliado devido a facilidade de coleta dos dados sem se considerar o grau em

que é representativa da população de interesse. A seleção da amostra é feita de forma intencional ou consecutiva.

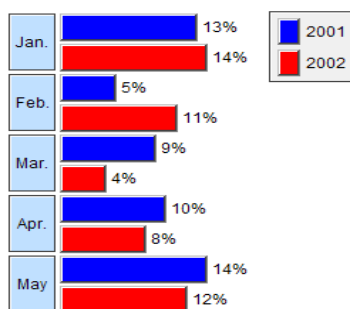
- AMOSTRAGEM PROBABILÍSTICA: implica um sorteio com regras bem determinadas, cuja realização só será possível se a população for finita e totalmente acessível. A seleção da amostra é realizada por sorteio.

8) GRÁFICOS:

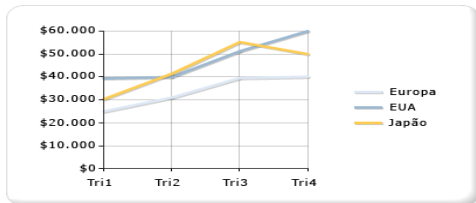
- COLUNAS: é composto por duas linhas ou eixos, pinto um vertical e outro horizontal.



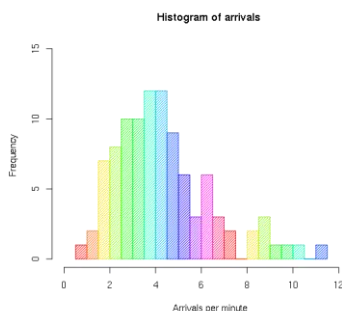
- BARRAS: são muito usados para comparar quantidades.



- LINHAS: é composto por dois eixos, um vertical e outro horizontal, e por uma linha que mostra a evolução de um fenômeno ou processo.



- **HISTOGRAMAS:** parecido com o gráfico de colunas em vários aspectos, pois sua construção é praticamente igual, mas seu cálculo é feito pela área do retângulo representado no gráfico.



9) TIPOS DE DISTRIBUIÇÕES:

- **SIMÉTRICA:** são dados concentrados no centro na distribuição.

Ex: Proteínas.

- **ASSIMÉTRICA POSITIVA:** são dados concentrados no lado esquerdo da distribuição. Ex: Triglicerídeos.

- **ASSIMÉTRICA NEGATIVA:** são dados concentrados no lado direito da distribuição. Ex: hemoglobina.

10) ESTATÍSTICAS DESCRITIVAS:

- **MÉDIA:** é a medida de tendência central que indica o valor provável de uma variável também chamado de valor esperado. Ela é afetada por valores extremos e não deve ser utilizada em distribuições assimétricas.

- **MEDIANA:** é um valor que tenha igual quantidade de valores menores e maiores do que ele e não é afetada pelos valores extremos. Deve ser utilizada em distribuições assimétricas para substituir a média.

- **MODA:** é o valor mais freqüente de uma série de valores e aparece como um pico de freqüência.

- **AMPLITUDE DE VARIAÇÃO:** é a diferença entre os valores extremos.

- **DESVIO PADRÃO:** como medida descritiva de dispersão e possui número diferente da variável que representam os desvios em torno da média. Se o desvio padrão for maior do que a metade da média, isso indica grande variabilidade e a distribuição passa a ser assimétrica.

- **AMPLITUDE INTERQUARTÍLICA:** são valores que dividem uma série ordenada de dados em quatro grupos de 25% que representam os percentis em 100 partes. Quando um conjunto de dados de distribuição assimétrica, a distância entre quartis representa melhor a variação do que a amplitude ou o desvio padrão, porque não é afetada pelos valores extremos.

DICA:

- *Distribuição simétrica: média e desvio padrão*

- *Distribuição assimétrica: mediana e amplitude interquartílica*

11) TESTES DE HIPÓTESES:

Método para verificar se os dados são compatíveis com alguma hipótese, podendo muitas vezes sugerir a não-validade de uma hipótese. O teste de hipóteses é um procedimento estatístico baseado na análise de uma amostra, através da teoria de probabilidades, usado para avaliar determinados parâmetros que são desconhecidos numa população.

- **NÍVEL DE SIGNIFICÂNCIA(&):** a probabilidade de rejeitar a hipótese nula quando ela é efetivamente verdadeira (ERRO). Também chamado de confiança que gira nos 5%.

- **HIPÓTESE NULA(h_0):** é a hipótese que traduz a ausência do efeito que se quer verificar.

- **HIPÓTESE ALTERNADA(h_1):** é a hipótese que o investigador quer verificar.

- **ERRO PADRÃO:** erro em uma série de medidas.

Erro tipo I: (falso- positivo) probabilidade de rejeitar a hipótese nula quando esta é verdadeira. O erro existe em afirmar uma diferença quando ela não existe. quando esta é falsa. O erro existe em afirmar uma igualdade quando o correto seria afirmar uma diferença.

Poder do teste: probabilidade de rejeitar a hipótese nula quando esta é falsa.

PODER= 100% - Beta = deve dar no mínimo de 80%

Valor (p) nível descritivo amostral: a hipótese nula funciona como um gatilho, o qual pode ser derrubada como falsa por um teste estatístico. Em uma estatística utiliza-se para sintetizar o resultado de um [teste de hipóteses](#) e define como a probabilidade de se obter uma estatística de teste igual ou mais extrema quanto aquela observada em uma amostra, assumindo verdadeira a hipótese nula. É comparado com o nível de significância.

Por exemplo, um nível-p de 0,05 (1/20) indica que há 5% de probabilidade de que a relação entre as variáveis, encontrada na amostra. Em outras palavras, assumindo que não haja relação entre aquelas variáveis na população, e o experimento de interesse seja repetido várias vezes, poderia-se esperar que em aproximadamente 20 realizações do experimento haveria apenas uma em que a relação entre as variáveis em questão seria igual ou mais forte do que a que foi observada naquela amostra anterior. Em muitas áreas de pesquisa, o nível-p de 0,05 é costumeiramente tratado como um "limite aceitável" de erro.

Dica: SE $P < \alpha$, O RESULTADO É SIGNIFICATIVO

SE $P > \alpha$, O RESULTADO **NÃO** É SIGNIFICATIVO

- TESTES DE HIPÓTESES PARA DUAS MÉDIAS:

AMOSTRAS INDEPENDENTES: são amostras que não tem relação entre si. Também poderemos dizer que variáveis que são manipuladas enquanto que outras dependentes são apenas medidas ou registradas. Ex: Teste t de student em amostras independentes

AMOSTRAS PAREADAS: essas possuem alguma relação de dependência entre si e são consideradas em planejamentos nos quais são realizadas duas medidas na mesma unidade amostral, ou seja, dados pareados, onde a unidade é o seu próprio controle.

Ex: Teste t de student em amostras pareadas.

- TESTES DE HIPÓTESES PARA TRÊS MÉDIAS OU MAIS:

AMOSTRAS INDEPENDENTES: para testes de comparação de médias com distribuição normal e variâncias iguais. Utiliza-se a Análise de Variância(ANOVA); One- Way; Post-Hoc: Tukey ou Bonferroni

AMOSTRAS PAREADAS: em comparação as médias com distribuições normais e variâncias iguais. Neste caso utiliza-se Análise de Variância (ANOVA) para medidas repetidas; Post-Hoc: Bonferroni

CORRELAÇÃO LINEAR DE PEARSON: mede o grau da correlação (e a direção dessa correlação - se positiva ou negativa) entre duas variáveis quantitativas de escala métrica (intervalar ou de rácio/razão).

Ex: utilizado na interpretação de pacientes no pré e pós operatório.

- TESTES NÃO-PARAMÉTRICOS: utilizados quando os dados são assimétricos ou existe uma grande diferença na variabilidade entre os grupos.

Exemplos de testes: U de Mann – Whitney ; T de Wilcoxon; H de Kruskai-Wallis; Friedman; Correlação de Spearman.

- TESTES PARAMÉTRICOS: (também chamados testes t) Exigem que a amostra tenha uma distribuição normal, possuam um desvio padrão baixo, especialmente deve ter uma dimensão inferior a 30 e são simétricos.

Exemplos de testes: T- student independente e pareado; ANOVA one- way; ANOVA para medidas repetidas; Correlação de Pearson.

- TESTES PARA PROPORÇÕES:

Amostras Independentes: Teste Qui-quadrado de Pearson (muito utilizado na área médica).

Amostras pareadas: Teste Qui-quadrado de McNemar.

- TESTES PARA PROPORÇÕES:

RAZÃO DE PREVALÊNCIA(RP): trabalha com prevalências, isso resulta em estudos transversais.

ODDS RATIO(OR): trabalha com a chance de exposição nos casos controles e indicado em estudos de caso-controle.

RISCO RELATIVO(RR): trabalha com incidências, isso resulta em estudos de coorte e ensaio clínico randomizado.

INTERVALO DE CONFIANÇA(IC): são usados para indicar a confiabilidade de uma estimativa de uma pesquisa que resulte IC pequeno é mais confiável do que uma que resulte IC maior. IC varia entre 0 a 1.

- FORMATAÇÃO ESTATÍSTICA DO BANCO DE DADOS :

- Cada indivíduo é uma linha no banco de dados;
- Cada coluna é uma variável;
- O banco de dados representa uma planilha somente;
- Variáveis quantitativas são inseridos os próprios valores;
- Variáveis qualitativas são inseridos atribuições e classificações;
- A primeira linha deve conter os nomes das variáveis;
- Sempre usar vírgulas ao invés dos pontos;
- Dar preferência a dados quantitativos;
- Sempre utilizar números;
- Cada questionário deve ter um número de referência;
- Não colocar a unidade da variável dentro do banco de dados;
- Cada variável deve ter um nome.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1- Hulley, Stephen B. Delineando a Pesquisa Clínica - Uma Abordagem Epidemiológica - 3ª Ed; Artmed, Porto Alegre, 2001.
- 2- Callegari-Jacques, Sídia M. (2003) Bioestatística: princípios e aplicações. Artmed. Porto Alegre.
- 3- Morettin, P.A. & Bussab, W.O. Métodos Quantitativos. 4ª Ed., São Paulo, Atual Editora Ltda., 1991. 321 p. (Métodos Quantitativos, Vol. 4).
- 4- Costa Neto, P.L.O. Estatística. 7ª Ed., São Paulo, Editora Blucher Ltda., 1987. 264 p.
- 5- Hoel, P.G. Estatística Elementar. Rio de Janeiro, Editora Atlas, 1989.
- 6- Dixon & Massey. Introduction to Statistical Analysis. McGraw Hill, 1969.
- 7- Oliveira L.A . - Probabilidade E Estatística Para Engenheiros - Apostila, Des-UFSCar
- 8- Oliveira L. A . - Estatística Aplicada À Educação - Apostila, Des-UFSCar
- 9- Barros Neto, B.; Scarminio, I.S., Bruns, R.E. - Planejamento e Otimização de Experimentos - Editora UNICAMP - 2ª. edição