



“...promoting safer aquatic environments through evidence-based research...”



8. MATERIAL SUPLEMENTAR SOBRE ESTATÍSTICAS (apenas como curiosidade).

- **Prevalência (*prevalence*):** número total de casos existentes numa determinada população num determinado momento temporal. Quantifica o risco do individuo de adquirir a doença naquela população, naquele determinado tempo. $P = \text{número de casos} / \text{população total} = \%$
- **Incidência (*incidence*):** número de novos casos surgidos numa determinada população num determinado intervalo de tempo.
- **Incidência cumulativa (*cumulative incidence (CI)*)** = proporção de pessoas que adoecem durante um período de tempo específico. $CI = \text{N total de novos casos por um período} / \text{população em risco}$
- **Densidade de Incidência (*incidence Density (ID)*)** – N novos casos num período de tempo / total população – tempo de observação.
- **Taxa de incidência (*incidence rate (IR)*)** (IR) = N total de novos casos por um período / população em risco neste tempo avaliado. Ex; A cidade X tinha 100 casos de doença de chagas em um ano e no ano seguinte passou a ter 120 casos. Surgiram 20 casos novos, então a incidência é 20 casos por 100000 habitantes.
 - A observação de 16 pessoas por 1 ano, ou 8 por 2 anos, ou 32 por 6 meses, fornecem o mesmo resultado.
 - 6 CASOS POR 100.000 por ano = 6/105/ano
 - 100(102); 1.000(103); 10.000(104); 100.000(105); 1.000.000(106);
- **Taxa de incidência acumulativa** = CI expostos (Cie) / CI não expostos (Cio)

TESTES PARAMÉTRICOS E NÃO PARAMÉTRICOS

Para testar as hipóteses de saber se utilizamos os testes paramétricos ou não paramétricos.

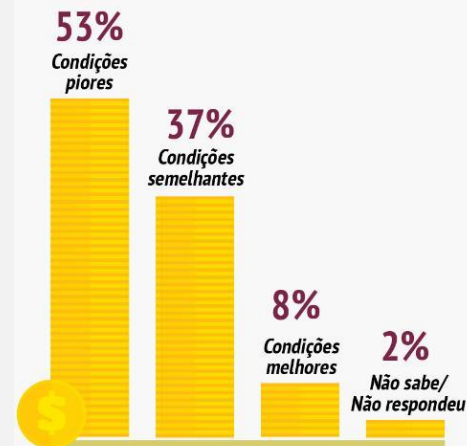
Testes paramétricos

Os testes paramétricos baseiam-se em medidas intervalares da variável dependente (um parâmetro ou característica quantitativa de uma população) e a utilização deste tipo de testes exige que sejam cumpridos três requisitos:

<p>1. Distribuição normal</p> <p>Os testes paramétricos são válidos quando aplicados a dados que obedecem a uma distribuição normal – uma distribuição normal é aquela que é perfeitamente simétrica à volta da média; tem a forma de um sino</p>	
<p>2. Homogeneidade dos dados</p> <p>No entanto, existem distribuições normais assimétricas, desviadas à direita ou à esquerda. Uma distribuição normal é aquela cuja análise estatística pode ser feita com dados da própria amostra, como a média, moda, mediana e desvio padrão. Quanto maior for o enviesamento das distribuições, menor será a validade do teste paramétrico.</p>	

3. Variáveis intervalares e contínuas.

Os testes paramétricos, tal como estão concebidos, podem aplicar-se apenas em dados (medidas relativas à variável dependente) que constituem uma escala de intervalos, ou seja, têm entre si intervalos contínuos e iguais.



Não Paramétricos

Os testes não paramétricos requerem menos pressupostos para as distribuições e têm maior relevância nas ciências sociais aplicadas. Baseiam-se em dados ordinais e nominais e são muito úteis para:

- Análise de testes de hipóteses;
- Análise de amostras grandes, em que os pressupostos paramétricos não se verifiquem, ou que não se tenha certeza de que sejam provenientes de população com distribuição normal.
- Amostras muito pequenas,
- Investigações que envolvam hipóteses cujos processos de medida sejam ordinais.

Os testes não paramétricos não são tão fidedignos como os testes paramétricos. São adequados para apoiar a tomada de decisão em situações nas quais não seja atendido algum dos requisitos para a aplicação dos testes paramétricos, como o teste Z, o teste T, e o teste F de análise de variância – ANOVA, que dependem: (i) da condição de a amostra ter sido extraída de uma população distribuída de acordo com distribuição normal (de Gauss); (ii) da escala de medida da variável aleatória ser contínua; e (iii) do tamanho da amostra ser maior do que 30 observações.

ASSOCIAÇÃO ESTATÍSTICA E RELAÇÃO CAUSA-EFEITO

ASSOCIAÇÃO ESTATÍSTICA

Todo estudo científico necessita uma avaliação da validade da associação feita através de VALIDAÇÃO ESTATÍSTICA. Existe 3 possibilidades:

1. ACASO (PLAY OF CHANCE) – avalia a possibilidade de a amostra estudada refletir ou não a realidade da população de interesse. Quanto maior a amostra utilizada menor a variabilidade e maior a confiança da inferência feita. Esta possibilidade é mensurada através de testes de significância estatística. $P \leq 0.05$, significa: Chance menor de 5% de ter ocorrido acaso no resultado ou 1 para cada 20 casos. Neste caso o intervalo de confiança também é analisado. Quanto mais estreito o intervalo, maior a confiança e maior a necessidade de amostra. Quanto maior o intervalo de confiança, maior a variabilidade do resultado e menor necessidade de amostra. A precisão é a medida de probabilidade do acaso (intervalo de confiança).

2. VIÉS (BIAS) (erro sistemático) – Em estudos descritivos, quanto menor o número de considerações e avaliação de erros menor o viés. Nos estudo analítico quanto maior a diferença entre os 2 grupos comparados (pacientes e controle) maior o viés.

Viés de SELEÇÃO – Os grupos comparados devem ser o mais similar possível para reduzir o viés.

Viés de OBSERVAÇÃO

- Do entrevistador – A coleta de dados deve ser o mais imparcial e cega possível
- Entrevistado – Quando o paciente exagera ou atenua suas respostas segundo seus interesses.

Viés de Exclusão – pacientes não recebem a intervenção completa, ou a recebem de forma errada ou no paciente errado.

3. Viés de CONFUSÃO (confundimento – Confounding) – Quando existe ou podem existir fatores não considerados como causa daquele efeito que foram considerados ou não e podem confundir o resultado final.

RELAÇÃO CAUSA-EFEITO

Busca caracterizar como causal uma associação entre uma exposição e uma doença ou condição de saúde. São nove critérios, onde quanto mais critérios forem preenchidos, maior a chance de esta associação ser de "causa e efeito":

1. Temporalidade (Time sequence) – o fator causal deve preceder a doença – único fator absolutamente imprescindível.
2. Força da associação (strength of the association) – questiona a relação causa-efeito quanto a sua força. Quanto mais forte uma associação, mais provável que seja causal. A força da associação é medida pelo risco relativo(RR) ou pelo odds ratio(OR).
3. Gradiente biológico (efeito dose-resposta) (Dose-response relationship) – A dose ou a quantidade da substancia ou do contato relaciona diretamente e progressivamente com o efeito. Quanto maior o contato(causa) maior o efeito e quanto maior efeito poderá aumentar Tb a causa. Ex. O aumento da agricultura aumentou a população e o aumento da população aumenta ainda mais a agricultura. No entanto no caso da causa ser especifica (especificidade) pode não haver aumento da causa com o aumento do efeito.
4. Consistência com outros estudos (consistency with other investigations) – A existência de outros trabalhos com diferentes métodos e população que revelam os mesmos resultados reafirmam a relação de causa-efeito.
5. Especificidade: exposição específica de um efeito causa a doença. Difícil de ser comprovado, por ser usualmente multifatorial. Therefore, it is necessary to examine specific causal relationships within a larger systemic perspective.
6. Plausibilidade biológica (Plausibility): A associação é plausível com o nosso entendimento dos processos patológicos existentes.
7. Credibilidade biológica Consideration of Alternate Explanations ((biologic credibility) – Quando existe um fator biológico já comprovado que indiretamente influencia como causa no efeito. Ex: o uso do álcool aumenta o HDL que então reduziria a coronariopatia.
8. Evidências experimentais: Mudanças na exposição mudam o padrão da doença
9. Coerência (Coherence): os achados devem seguir o paradigma da ciência atual

Análise estatística em distribuição normal – descrição paramétrica (testes paramétricos)

É aquela cuja análise estatística pode ser feita com dados da própria amostra, como a média, moda e mediana. Estas são medidas de tendência central ou Medidas de Posição que indicam a localização dos dados.

Medidas de dispersão - As medidas mais comuns de variabilidade para dados quantitativos são a variância; a sua raiz quadrada, o desvio padrão. A amplitude total, a distância interquartílica e o desvio absoluto são mais alguns exemplos de medidas de dispersão.

Numa amostra que não tem uma distribuição normal não é possível calcular o desvio padrão, por exemplo.

Média - Em estatística a média é o valor que aponta para onde mais se concentram os dados de uma distribuição.

Mediana - Em teoria da probabilidade e em estatística, a mediana é uma medida de tendência central, um número que caracteriza as observações de uma determinada variável de tal forma que este número (a mediana) de um grupo de dados ordenados separa a metade inferior da amostra, população ou distribuição de probabilidade, da metade superior. Mais concretamente, 1/2 da população terá valores inferiores ou iguais à mediana e 1/2 da população terá valores superiores ou iguais à mediana.

A mediana pode ser calculada para um conjunto de observações ou para funções de distribuição de probabilidade.

Moda - Em estatística descritiva, a moda é o valor que detém o maior número de observações, ou seja, o valor ou valores mais freqüentes, ou ainda "o valor que ocorre com maior freqüência num conjunto de dados, isto é, o valor mais comum". O termo moda foi utilizado primeiramente em 1895 por Karl Pearson, sob influência do termo moda referindo-se ao uso popular com o significado de objeto que se está usando muito no tempo presente. A moda não é necessariamente única, ao contrário da média ou da mediana. É especialmente útil quando os valores ou observações não são numéricos, uma vez que a média e a mediana podem não ser bem definidas.

Confiabilidade do Estudo

Validade externa (generalidades) – resultado tem de ser generalizado para outras circunstâncias (critério de inclusão e exclusão)

Validade interna (entra o viés) – viés de seleção, informação, confundimento.

Precisão (P) – Probabilidade e intervalo de confiança

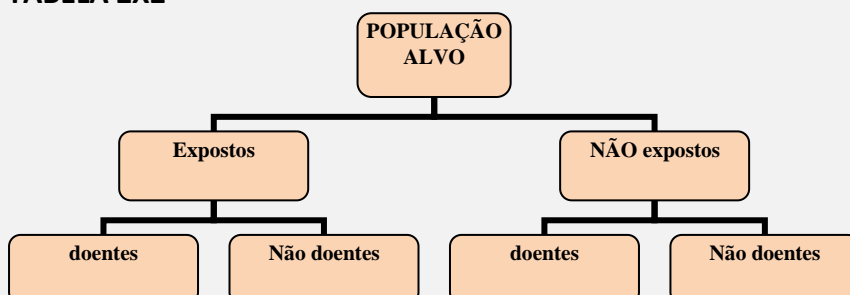
Medidas de efeito – teste de significância estatística

Medidas de efeito são medidas construídas através de razões ou de diferenças

RISCO

Probabilidade de ocorrência de certo desfecho, varia entre 0 e 1 e pode ser transformado em percentual ao se multiplicar por 100. É preciso cautela ao interpretar o aumento do risco, pois nem sempre significa algo indesejado.

TABELA 2X2



RISCO RELATIVO (RR) (risk ratio ou relative risk)

Indica a possibilidade de desenvolver a doença no grupo exposto, relativo ao grupo não exposto – Uso em Coorte (Cohort). É importante especificar o tempo calculado como de exposição, já que o número de casos varia conforme o tempo.

$RR = \frac{\text{incidência da doença no grupo exposto}(I_e)}{\text{incidência de doença no grupo não exposto}(I_o)}$
 C_e (incidência cumulativa expostos); C_o (incidência cumulativa não expostos)

Valores apurados:

- = 1 = Risco entre expostos e não expostos
- < 1 = Risco maior entre os não expostos
- > 1 = Risco maior entre os expostos

Ex: RR de 1.4 = 40% mais de chance de adquirir a doença entre os expostos.

Redução do risco relativo (RRR) ou eficácia = $(1-RR) \times 100 = X\%$

Excesso relativo de risco (ERR) = $(RR-1) \times 100 = X\%$

REDUÇÃO ABSOLUTA DE RISCO (RAR)

É a redução em termos absolutos do risco no grupo que sofreu a intervenção de interesse, em relação ao grupo controle. $RAR = [R(T) - R(C)] \times 100$

Number need to treat (número necessário para tratar) NNT

$NNT = 1 - RAR$, ou número de pacientes que se precisa tratar para que 1 se beneficie.

NNH = Number need to Harm = número de pacientes que se precisa tratar para ocorrer um dano.

RISCO ATRIBUÍDO (ATTRIBUTABLE RISK)

É a diferença de incidência entre os grupos expostos (I_e) e o grupo não expostos (I_o)

$AR = I_e - I_o$

ODDS

Não é uma probabilidade, pois o denominador não inclui o numerador e não pode ser expresso em %. Se em uma população de 1000 pessoas, 400 adoece, o risco é de 0,40 (400/1000) ou 40%, enquanto o ODDS, é de 400/600 ou 2:3.

ODDS RATIO (OR)

Utilizado para estudo caso-controle, já que os pacientes são selecionados através da doença, e é difícil calcular os casos expostos e não expostos. O OR mostra o RR nos estudos caso-controle.

$OR = \frac{a/c}{b/d} = \frac{ad}{bc}$

Quando a doença é rara, a proporção de casos expostos e não expostos é baixo, o número total de expostos ($a+b$), pode ser aproximado utilizando-se o número de controle expostos (b) e o total de não expostos ($c+d$) pelo número de não expostos controle (d). Assim o RR é representado pelo OR.

ABSOLUTE ODDS RATIO (AOR)**POPULATION ATTRIBUTABLE RISK (PAR)**

O risco relativo (RR) julga se uma associação observada é válida ou casual. O Risco Atribuído (AR) nos dá a medida do impacto na saúde pública, assumindo que a associação tem causa-efeito.

Exemplo: 20 anos de observação, fumando 20 cigarros por dia, mostra 14 vezes mais câncer e 1.6 mais doença coronária em fumantes. Portanto o cigarro é um fator causal muito mais forte para o câncer de pulmão do que para doença coronária. Entretanto parar de fumar determinará em prevenção muito maior de doença coronária (256/100.000/ano) do que no câncer de pulmão (130/100.000/ano).

TABELAS “r-by-c” table

Utilizam-se quando existem vários níveis de exposição ou diferentes categorias de algumas características da população estudada.

Exposição	Doença		Total
	Sim(+)	Não(-)	
Sim(+)	a	b	a + b
<5 anos			
5 a 10 anos			
10 a 15 anos			
Não(-)	c	d	c + d
	a + c	b+d	a + b + c + d

TAXA DE MORTALIDADE PROPORCIONAL (proportional mortality ratio)

$$\text{PMR} = \frac{\text{mortes por causa especifica(expostos)}}{\text{mortes por causa especifica sobre a população geral}} \times 100 = \%$$

Calcula a mortalidade para um doença entre os expostos a um fator comparado a população total sob risco

TESTE DE SIGNIFICANCIA ESTATÍSTICA

A evidencia é avaliada contra a hipótese nula de que não há diferença entre os tratamentos, através do P-valor (nível de significância), ou seja a probabilidade de encontrar aquela diferença, quando ela não existe.

P-valor: é a probabilidade de achar resultado + ao acaso (ou seja falso +) e aceita-se que deve ser menor de 5% ($P < 0.05$).

Nível de significância $< 5\%$ (0,05) é forte evidencia de que haja uma associação e não magnitude.

Eficácia = magnitude da associação, quanto maior, maior a associação (grande eficácia).

Intervalo de confiança = intervalo com probabilidade de incluir o valor real que se esta externando. Se usa geral/ 95% de intervalo de confiança.

Precisão – quanto menor o intervalo de confiança, maior a precisão.

Estudo inconclusivo – o intervalo de confiança contem o valor nulo d efeito (sem significância estatística) ou $RRR=1$ ou $RAR=0$.

Poder – é a probabilidade de o estudo identificar uma diferença entre os tratamentos (efeito) quando esta diferença é real. Exemplo: um homem entre 55 e 64 anos tem taxa de mortalidade por doença coronária de 2/1000 por ano em 2001. Ou, a cada 100, em média 2 morrerão de doença coronária em 1 ano. Ou, um risco de 99,8% de que não morra de doença coronária no próximo ano.