

4. AMOSTRAGEM, COLETA E ORGANIZAÇÃO DOS DADOS

1. **Regra Básica: evitar viés (desvios): Amostra deve ser absolutamente aleatória**
 - 1.1. Cada indivíduo da população deve ter a mesma chance de ser escolhido
 - 1.2. Amostras auto selecionadas não têm validade científica
2. **Definição da População-Alvo: Todo o grupo de indivíduos que você quer estudar**
 - 2.1. Deve ser bem definida para que a amostra possa ser bem definida
 - 2.2. Opinião dos usuários de transporte coletivo em Ribeirão Preto: todos os usuários
3. **Definição da Amostra: deve ser representativa da população-alvo**
 - 3.1. Porcentagem da população, segundo erro amostral suportável 2% a 5%
 - 3.2. Erro amostral: fórmulas aproximadas e fórmulas mais precisas
 - 3.2.1. Fórmulas aproximadas: tamanho da população e erro suportável
 - 3.2.2. Fórmulas mais precisas: tamanho da população e variância na população
 - 3.2.2.1. Necessário estudo-piloto antes de definir a amostra

$$E = \frac{1}{\sqrt{n}}$$

$$n = \frac{N * \frac{1}{E^2}}{N + \frac{1}{E^2}}$$

n: tamanho da amostra
N: tamanho da população
E: erro amostral (decimal)

$$n = \left[\frac{Z_{\alpha/2} * \sigma}{E} \right]^2$$

n: tamanho da amostra
 $Z_{\alpha/2}$: Valor z para intervalo de confiança
 σ : desvio padrão populacional
E: erro amostral (decimal)

$$E = \sqrt{\frac{N - n}{N * n}}$$

- 3.3. Amostra deve ser escolhida aleatoriamente dentro da população-alvo
 - 3.3.1. Amostras não aleatórias trazem viés (distorção, tendências) nos resultados
 - 3.3.2. Amostragem por telefone celular: pessoas com mais de 1 celular (aceitável?)
 - 3.3.3. Pesquisas em websites: amostra acessa internet, aquele site, decidem responder
 - 3.3.4. Amostra deve ser analisada buscando-se possíveis vieses
 - 3.3.4.1. Estudo de hábitos de estudo pesquisando frequentadores da biblioteca
 - 3.3.4.2. Estudo de um novo tipo de cirurgia de catarata: ética x voluntários
 - 3.3.4.3. Solução: grupo de tratamento e grupo de controle similares (duplo cego)
- 3.4. Tipos de Amostragem
 - 3.4.1. Amostragem aleatória simples: escolha aleatória dentro da população
 - 3.4.2. Amostragem aleatória estratificada (cluster): separação em estratos uniformes e dentro de cada estrato a escolha é aleatória (n pessoas em cada extrato)
 - 3.4.3. Amostragem por Conglomerados: semelhante à estratificada simples, mas proporcional ao número de elementos em cada conglomerado (bairros)
 - 3.4.4. Amostragem sistemática: cada 10º elemento em uma linha de produção
 - 3.4.5. Amostragem a esmo (parafusos na caixa),
 - 3.4.6. Amostragem intencional e amostragem por voluntários (auto selecionada);
 - 3.4.7. Amostragem por Conveniência (não probabilística): sem rigor estatístico. Seleciona os elementos a que tem acesso, admitindo que estes possam representar a população.
- 3.5. Definida a amostra, as “não respostas” devem ser minimizadas (prêmios, brindes, etc.)
 - 3.5.1. Prêmios ou sorteios ou brindes podem causar viés – atraem um tipo de pessoas
 - 3.5.2. Importância do anonimato e confidencialidade das respostas
- 3.6. Avaliação crítica: Amostra planejada x Amostra efetiva
 - 3.6.1. Amostra efetivamente consultada deve ser, no mínimo, 70% da amostra planejada
 - 3.6.2. Composição da amostra efetiva deve ser similar à da amostra planejada
 - 3.6.3. Importância da busca de outros possíveis desvios entre planejamento e realidade
 - 3.6.4. Reportagem do New York Times afirma que consumo de maconha é muito mais efetivo no alívio dos efeitos da quimioterapia do que outras drogas:
 - 3.6.4.1. 29 pacientes, 15 com maconha, 14 com placebo (12 terminaram estudo)

4. Importância do Efeito-Placebo

- 4.1. Efeitos fisiológicos causados pela sugestão (psicológicos)
- 4.2. Doenças psicossomáticas: 80% das doenças
- 4.3. Situações éticas onde não se usa placebo, mas tratamentos convencionais

5. Questionários

- 5.1. Elaboração das questões e sua ordem no questionário são fundamentais
 - 5.1.1. Objetividade na elaboração do questionário: foco no objetivo do estudo
 - 5.1.2. Você concorda que o presidente tenha que viajar ao exterior representando o país?
 - 5.1.3. Você apoia a pena de prisão perpétua para sequestradores (após reportagem)
- 5.2. Pessoas que aplicam o questionário devem ser bem treinadas
 - 5.2.1. Abordagem do entrevistado é fundamental
 - 5.2.2. Dúvidas sobre as perguntas devem ser respondidas da mesma forma por todos
 - 5.2.3. Em grandes levantamentos, estudos-piloto em pequena escala são importantes

6. Identificação de variáveis de confusão

- 6.1. São variáveis não incluídas no estudo, mas que influenciam fortemente os resultados
- 6.2. Ex: Pessoas que tomam vitamina C diariamente têm menos gripe:
 - 6.2.1. Estudo observacional, quem toma vitamina C pode ter mais preocupação com a saúde e se cuidar melhor, por isso ter menos gripe
 - 6.2.2. Solução: pessoas com mesmo nível de preocupação com a saúde, divididas aleatoriamente, consome vitamina C, consome Placebo, não consome nada

7. Qualidade dos Dados

- 7.1. Confiabilidade: resultados são reprodutíveis em medidas subsequentes
 - 7.1.1. Medidas de pressão sanguínea, açúcar no sangue, temperatura corporal, etc.
- 7.2. Dados não enviesados: sem favoritismo sistemático em certas respostas individuais.
 - 7.2.1. Erros no instrumento de medida, amostras viciadas, respostas induzidas
 - 7.2.2. Vieses devem ser prevenidos – são de difícil correção depois de colhidos os dados
- 7.3. Validade: dados representam o que eles têm a intenção de medir
 - 7.3.1. Prevalência de crimes em um bairro/cidade não é o número absoluto de crimes no bairro/cidade, mas o número de crimes por habitante do bairro/cidade

8. Pontos a observar em um estudo:

- 8.1. Gráficos tendenciosos (equivocados)
- 8.2. Dados enviesados (tendenciosos)
- 8.3. Margem de erro não reportada (maior do que se pensa)
- 8.4. Tamanho da Amostra (pode ser insuficiente, levando a erro amostral elevado)
- 8.5. Amostra não aleatória (levam a tendências nos resultados)
- 8.6. Dados amostrais perdidos (nem toda a amostra planejada foi analisada)
- 8.7. Correlações equivocadas
 - 8.7.1. correlação: dados quantitativos (peso e altura) – não implica causa-efeito
 - 8.7.2. associação: dados qualitativos (votos femininos para o candidato)
 - 8.7.3. força de uma correlação – o R^2 não relatado
- 8.8. Variáveis de confusão (não previstas no estudo, mas com grande influência)
 - 8.8.1. Frequentadores de academia tem melhor saúde (estilo de vida: muitas variáveis)
- 8.9. Resultados reportados parcialmente
 - 8.9.1. Resultados estatisticamente significativos: em 1 dos 50 testes realizados?
 - 8.9.2. Resultados individuais – Artista X perde 10 kg em uma semana com a dieta Y

9. Organização dos Dados: tabulação em dois tipos

- 9.1. Cada linha é um registro (entrevistado), cada coluna uma pergunta (variável)
- 9.2. Cada linha é uma pergunta (variável), cada coluna um registro (entrevistado)

