

Resumo 01 – Planejamento de uma Pesquisa

A necessidade de encontrar solução prática para determinado problema ou curiosidade científica é fator determinante para que o investigador se proponha a explorar e a planejar experimentos que tragam conhecimentos ou que venham a rejeitar ou ratificar antigas experiências.

1.1. Escolha do tema e definição do problema

O procedimento científico tem início com a escolha de um tema geral e a definição de um problema que mereça investigação.

O tema de uma pesquisa é o assunto que se deseja desenvolver e que contém um problema relevante, o qual representa uma dificuldade ainda não solucionada ou solucionada de maneira controvertida.

É necessário determinar com precisão o problema decorrente do assunto para que se formule uma explicação provisória (hipótese) que será avaliada criticamente mediante o delineamento de um experimento (experimentação). Esta propiciará a formulação de uma proposição final com base no teste da hipótese (que será rejeitada ou aceita).

Enunciar o problema implica determinar o objetivo central da indagação. Assim, enquanto o tema da pesquisa pode ser encarado como mais abrangente, a formulação do problema é mais específica, indicando de forma clara e precisa a dificuldade que se pretende resolver, limitando o seu campo e destacando suas características.

O problema consiste em um enunciado apresentado de forma clara e compreensível, cuja solução poderá ser conseguida com o planejamento de uma pesquisa experimental.

1.2. Formulação dos objetivos

Devem ser enunciados de forma bastante clara, já que as demais etapas da pesquisa tomam como base esses objetivos. A elaboração dos objetivos específicos deve ser feita de tal forma que forneça uma primeira indicação das características que são necessárias observar ou medir nos elementos pesquisados.

1.3. Formulação das hipóteses

Uma vez formulado o problema propõe-se uma explicação provável, isto é, uma hipótese. Tanto o problema como a hipótese são enunciados de relações entre fatos e fenômenos.

Em toda pesquisa, exceto naquelas meramente descritivas ou expositivas, cabe, do ponto de vista estatístico, a formulação de uma hipótese para sua comprovação ou rejeição.

A hipótese formulada como descrito acima é passível de teste mediante um experimento que poderia ser conduzido com dois grupos experimentais (bovinos europeus e bovinos indianos) e que proporcionaria elementos para sua rejeição ou não.

Uma afirmação pode ser negativa ou positiva. Da maneira como foi formulada no exemplo, a hipótese é uma afirmação positiva. A formulação da hipótese por meio de uma afirmação negativa seria: Bovinos de procedência indiana e europeia não apresentam diferença quanto à adaptação em altas temperaturas ambientais.

Esta maneira negativa de formular a hipótese será vista novamente quando se discutir hipótese estatística, pois estas são duplamente formuladas de modo negativo (hipótese nula) e de forma afirmativa (hipótese alternativa).

As hipóteses dirigem a investigação, indicando o que se deve colocar à prova. Não é possível avançar na pesquisa se após definir o problema que originou o estudo não se propuser uma explicação ou solução provisória para ele. A hipótese orienta a busca do entendimento e da ordenação dos fatos. As sugestões formuladas na hipótese podem ou não ser as soluções do problema. Mas, é para saber se o são, que a pesquisa é realizada. Sem a orientação fornecida pela hipótese não é possível saber o que observar e que experiência realizar.

1.4. Delineamento da pesquisa (desenho do estudo)

1.4.1. Estudos Observacionais (A observação é feita sem interferência do pesquisador)

- **Pesquisa levantamento ou survey:** Observam-se características dos elementos de uma certa população ou amostra, utilizando-se questionários ou entrevistas.
- **Estudo de casos:** Consiste na cuidadosa e minuciosa descrição, por um ou mais pesquisadores, do diagnóstico e evolução da doença de um pequeno número de indivíduos, geralmente menos de 10.
- **Estudo de Série de Casos:** Consiste na apresentação organizada, por um ou mais médicos, dos resultados dos seus preferidos, na descrição dos pacientes tratados e na divulgação de outras informações que julgam relevantes. (>10). Envolvem pacientes acompanhados durante um período de tempo relativamente curto. Em geral não possuem pacientes-controle.
- **Estudos de Caso-controle (retrospectivo):** Os casos são indivíduos selecionados com base em alguma doença ou algum desfecho clínico. Os controles são indivíduos sem a doença ou o desfecho clínico. A história de eventos anteriores tanto dos casos como dos controles são analisadas na tentativa de identificar uma característica ou fator de risco presente nas histórias dos casos, porém não nas dos controles. Pergunta: O que aconteceu???
- **Estudos de Coorte (prospectivo):** Uma coorte é um grupo de pessoas que têm algo em comum e que permanecem parte de um grupo durante certo período de tempo. Os participantes de estudos de coorte são escolhidos por alguma característica (as) definidora presumida como sendo precursora ou fator de risco de uma doença ou efeito sobre a saúde. Pergunta: O que irá acontecer???

1.4.2. Estudos Experimentais

O pesquisador “manipula” de forma planejada certas *variáveis independentes* ou *fatores* (A, B, C, ...) para verificar o efeito que esta manipulação provoca numa certa *variável dependente* ou *resposta*. É usado para resolver problemas bem específicos formulados sob a forma de hipóteses de causa e efeito. Gera dados suficientemente estruturados, devido o controle do pesquisador, para que se possa decidir, através de uma análise estatística, se uma hipótese previamente formulada pode ser aceita ou não.

1.4.3. Metanálise (“meta” significando ulterior ou mais organizada)

Utiliza informações publicadas de outros estudos e associa os resultados de modo a permitir uma conclusão geral. É semelhante aos trabalhos de revisão de literatura, porém inclui além deles uma avaliação quantitativa e um resumo dos achados. Pode-se fazer uma metanálise de estudos observacionais e experimentais, mas deve descrever separadamente os achados desses dois tipos de estudo.

2. Métodos de coleta de dados

2.1. Fontes primárias

Quando não há registro de informação. A informação é obtida diretamente dos indivíduos nos quais pretendemos realizar o estudo.

2.2. Fontes secundárias

Quando a informação de interesse está registrada, basta fazer bom uso delas.

2.3. Amostragem

Uma população ou universo de dados, no sentido geral, é um conjunto de elementos com pelo menos uma característica em comum que possa ser contada, medida, pesada ou ordenada de algum modo. Essa característica comum deve delimitar inequivocamente quais os elementos que pertencem à população e quais os que não pertencem.

Em qualquer estudo estatístico, tem-se sempre em mente pesquisar uma ou mais características dos elementos de alguma população. Os dados que são observados, na tentativa de tirar conclusões sobre o fenômeno de interesse, serão referentes a elementos dessa população, daí a importância de se caracterizar perfeitamente a população em estudo.

São exemplos de população estatística: os animais de um bairro, as indústrias e as lojas de uma cidade, os estudantes de uma universidades, os habitantes de uma cidade, etc.

Uma vez perfeitamente caracterizada a população, o passo seguinte é o levantamento de dados acerca da(s) característica(s) de interesse no estudo em questão. Grande parte das vezes, porém, não é conveniente, ou mesmo nem é possível, realizar o levantamento dos dados referentes a todos os elementos da população.

Uma amostra é um subconjunto, representativo ou não, da população em estudo, necessariamente finito, pois todos os seus elementos serão examinados para efeito da realização do estudo estatístico desejado. A representatividade é uma propriedade altamente desejada em estatística, que ocorre quando a amostra apresenta as mesmas características gerais da população da qual foi extraída.

Denomina-se Censo ou Recenseamento ao exame completo de todos os elementos da população e Amostragem, quando a análise é através de uma parte, representativa, da população.

É intuitivo que, quanto maior a amostra, mais precisas e confiáveis serão as conclusões sobre a população. Deste modo, os resultados mais perfeitos seriam obtidos através do Censo. Mas, na prática, nem sempre isso é possível, obrigando à utilização da Amostragem. O emprego de amostras pode ser feito de tal modo que se obtenham resultados confiáveis. A definição da forma do estudo populacional, censo ou amostragem, depende de uma série de fatores, como: *custo do levantamento de dados*, *o tempo necessário para realizá-lo*; *riscos de os próprios agentes interferirem nas condições normais da população*, especialmente se a população for muito grande. Na área médica, por exemplo, o estudo de epidemias pode exigir a imediata aplicação de vacinas para controlá-las; os pesquisadores podem tornar-se os principais agentes de propagação dos vírus das doenças contagiosas que estão tentando debelar.

A estatística amostral, além de avaliar as relações existentes entre uma população e as diversas amostras que dela se pode extrair, também se presta para indicar se as diferenças observadas entre os parâmetros amostrais obtidos de duas ou mais populações são casuais ou significativas.

Outro dado importante é que a estatística amostral é o elo de ligação entre a *Teoria das*

Probabilidades e a Inferência Estatística (primeira é o instrumento de medida da Segunda). Tanto na Teoria das Probabilidades quanto na Inferência Estatística se busca a essência das questões em estudo. Na Teoria das Probabilidades, buscam-se resultados gerais que posteriormente possam ser aplicados a casos particulares; na Inferência Estatística, buscam-se resultados particulares (extraídos das amostras das populações em estudo) que se apliquem a casos gerais (população em estudo). Enfim, as duas juntas evidenciam as relações entre as amostras e as populações de onde elas foram extraídas.

2.3.1. Seleção da Amostra

O que qualifica um método de amostragem em bom ou ruim é sua possibilidade ou não de definir a probabilidade de um elemento qualquer da população ser incluído ou não na amostra. Um método ruim, por exemplo, é aquele que privilegia a inclusão na amostra de elementos particulares numa proporção maior que aquela em que eles existem na população. Um método bom é aquele em que todos os elementos, individualmente considerados, têm a mesma chance de compor uma amostra em que as características proporcionais da população são respeitadas.

A escolha do melhor método de amostragem passa pela correta interpretação do problema em questão e pela forma como as informações serão extraídas dos elementos selecionados. Uma amostragem aleatória, onde pessoas são abordadas na rua, ao acaso, por um investigador desconhecido, não é adequada a temas subjetivos ou com significativa conotação pessoal e emocional. O que se constata, em tais casos, é um alto índice de recusa dos indivíduos em cooperar com o estudo. Nesses casos, ou eles não respondem às questões formuladas ou, quando fazem, não se pode garantir que as informações fornecidas são honestas e completas. O sucesso aí depende muito da habilidade do entrevistador em lidar com as questões propostas, de sua capacidade em transmitir confiança e da receptividade dos entrevistadores. Como tais “qualidades” são variáveis de difícil controle, não se pode contar com elas como garantia de correção e representatividade dos dados coletados.

Uma vez tendo optado pelo processo de amostragem para realizar o estudo, é preciso estabelecer a técnica de amostragem, isto é, o procedimento que será adotado para escolher os elementos que irão compor a amostra. É preciso garantir que a amostra ou as amostras que serão usadas sejam obtidas por processos adequados. Se erros grosseiros forem cometidos no momento de selecionar os elementos da amostra, todo o trabalho ficará comprometido e os resultados finais serão provavelmente bastante incorretos. Portanto, cuidado especial deve ser tomado quanto aos critérios que serão usados na seleção da amostra. É necessário garantir que a amostra seja representativa da população, isso significa que, a menos de pequenas discrepâncias inerentes a aleatoriedade sempre presente, em maior ou menor grau, no processo de amostragem, a amostra deve possuir as mesmas características básicas da população, no que diz respeito à(s) variável(is) em estudo.

Existem dois tipos de amostragem: *a probabilística e a não probabilística*.

A amostragem probabilística (ou objetiva) é aquela em que todos os elementos da população têm probabilidade conhecida, e diferente de zero, de pertencer à amostra. A amostragem não-probabilística (ou subjetiva) é o contrário: a escolha dos elementos da amostra é feita de forma não aleatória, justificadamente ou não. A amostragem probabilística fornece melhor garantia da representatividade da amostra em relação a população, razão pela qual deve ser preferida sempre que for possível e o custo a ela associado for admissível. Se a amostra não for representativa, as conclusões sobre a população, da qual foi extraída, não

terão validade, e nem será quantificável o erro possível entre o parâmetro real e o estimado com base nos seus dados. A utilização de uma amostragem probabilística garante a representatividade da amostra, pois o acaso será o único responsável por eventuais discrepâncias entre os elementos da população e da amostra. Isso porque, para a amostragem probabilística, existem procedimentos seguros que nos permitem inferir a partir da amostra extraída da população de interesse.

2.3.2. Amostragem Probabilística

- Amostragem Aleatória ou Casual Simples

Nesse tipo de amostragem pressupõem-se que todos os elementos da população têm a mesma probabilidade de ser incluídos na amostra. Essa técnica pode ser realizada através de um simples sorteio ou mediante o emprego da tabela de *números aleatórios*. O processo consiste na listagem de todos os elementos da população-alvo e uso dos números da tabela para a escolha dos componentes da amostra. A desvantagem desse processo é a exigência da enumeração de todos os elementos da população e sua identificação posterior, quando escolhidos para compor a amostra.

- Amostragem Sistemática

Neste processo, as amostras são retiradas periodicamente a partir de determinado elemento ou ponto de partida, como por exemplo, de hora em hora ou segundo uma regra.

A amostragem sistemática possui uma grande facilidade na determinação dos elementos da amostra, contudo, o perigo em adotá-la está na possibilidade da existência de ciclos de variação da variável de estudo, especialmente se o período desses ciclos coincidir com o período de retirada dos elementos da amostra. Por outro lado, se a ordem dos elementos na população não tiver qualquer relacionamento com a variável de interesse, então a amostragem sistemática terá efeitos equivalentes à amostragem aleatória, podendo ser utilizada sem restrições. Apesar de seu inconveniente, é um método bastante usado em pesquisa de opinião pública, onde os transeuntes são entrevistados de hora em hora ou após certa contagem

- Amostragem por Conglomerados

Nesse processo, a população que se deseja estudar é subdividida em grupos fisicamente próximos, chamados de conglomerados, sem levar em consideração a sua homogeneidade. Nos conglomerados são agregados os elementos populacionais com estreito contato físico, como casas, quarteirões, bairros, cidades, etc. A vantagem desse processo é que muitos elementos podem ser estudados por poucos pesquisadores, economizando custos e tempo.

- Amostragem Estratificada

Em algumas situações é de interesse subdividir a população em grupos aproximadamente homogêneos, independentemente deles serem encontrados próximos ou não. Nesse caso, a população que se deseja estudar é subdividida em grupos homogêneos, com características comuns, chamados de estratos. Nos estratos são agregados os elementos populacionais com mesmo sexo, ou raça, ou religião, ou escolaridade, etc. A vantagem desse processo é que pequenas amostras são suficientes para representá-lo.

2.3.3. Amostragem Com e Sem Reposição

Na *amostragem com reposição* o elemento extraído é devolvido à população após anotadas suas características de interesse ao estudo, enquanto que, na *amostragem sem reposição* elemento extraído não é devolvido à população após anotadas suas características de interesse ao estudo. A diferença básica

entre um tipo e outro é a possibilidade de um elemento ser ou não considerado mais de uma vez na amostra que esta sendo produzida

A amostragem com reposição pode ser usada para pequenas populações e sem reposição para grandes populações. A não reposição de um elemento à população altera progressivamente as probabilidades de extração dos itens seguintes e anula a probabilidade de um mesmo elemento ser extraído mais de uma vez. Enquanto que nas grandes populações, essa alteração pode ser desprezada, nas pequenas, essa alteração torna-se cada vez mais sensível, especialmente nos casos em que a amostragem tende a esgotar a população.

Algumas situações justificam a realização de um ou outro tipo de amostragem. Nos chamados testes destrutivos, o que se pretende, por exemplo, é testar a eficácia de certa medicamento, razão pela qual deve-se aplicar o medicamento em um paciente, sendo assim impossível repor o elemento amostrado na população. Outro exemplo é quando o estudo amostral envolve altos custos (do material ou do próprio teste).

2.3.4. Número de Amostras

Sendo **N** o número de elementos da população e **n** o número de elementos da amostra, o número, **k**, de amostras possíveis é dado por:

- se a amostragem for *com reposição*: $k = N^n$ amostras possíveis;
- se a amostragem for *sem reposição*: $k = \frac{N!}{n! \cdot (N - n)!}$ amostras distintas possíveis;
- se a amostragem for *sem reposição*: $k = \frac{N!}{(N - n)!}$ amostras duplicadas possíveis.

Nota: Amostras distintas são formadas por distintos elementos e amostras duplicadas são formadas pela permutação de dois elementos das amostras distintas.

2.3.5. Tamanho da amostra

N = número de elementos da população

n = número de elementos da amostra

n_0 = uma primeira aproximação para o tamanho da amostra

E_0 = erro amostral tolerável

$$n_0 = \frac{1}{E_0^2}$$

se **N** é muito grande ou desconhecido

$$n = n_0$$

se **N** não for muito grande e for conhecido

$$n = \frac{N \cdot n_0}{N + n_0}$$

3. Variáveis

Variável é uma determinada característica que, por algum motivo, é de interesse estudar. É dita aleatória devido seus possíveis valores dependerem unicamente do acaso. A escolha da variável de interesse dependerá, em cada caso, dos objetivos do estudo estatístico em questão.

A característica de interesse de estudo (variável) pode ser dividida em duas categorias: *qualitativas* e *quantitativas*.

As variáveis qualitativas apresentam como possíveis resultados uma qualidade (ou atributo) do indivíduo pesquisado. Dentre as variáveis qualitativas, existe, ainda, uma distinção entre dois tipos: variável qualitativa categórica ou nominal, para a qual não existe nenhuma ordenação nos possíveis resultados, e variável qualitativa ordinal, para a qual existe uma certa ordem nos possíveis resultados.

As variáveis quantitativas apresentam como possíveis resultados números resultantes de uma contagem ou mensuração. Dentre as variáveis qualitativas, existe, ainda, uma distinção entre dois tipos: variáveis quantitativas discretas, cujos possíveis valores formam um conjunto finito ou infinito enumerável de números e que resultam, freqüentemente, de uma contagem; e variáveis quantitativas contínuas, cujos possíveis valores formam um intervalo de números reais e que resultam, normalmente, de uma mensuração.