



ESTATÍSTICA E PESQUISA DE OPINIÃO

***Prof. Paulo Ricardo B. Guimarães
Departamento de Estatística -UFPR***

INTRODUÇÃO

→ A maioria das pessoas pensa que Estatística é uma simples coleção de números, tem a ver com gráficos e Censo Demográfico;

→ A origem da palavra Estatística tem a ver com uma coleção de informações populacionais e econômicas de interesse do Estado;

→ Na verdade, é muito mais que isso! É um conjunto de técnicas de análise de dados, cientificamente formuladas, aplicáveis a quase todas as áreas do conhecimento que nos auxiliam no processo de tomada de decisão; E isto hoje em dia é muito importante!

→ Hoje em dia as portas do mercado de trabalho estão muito mais abertas aos profissionais que, por exemplo, tem habilidades em línguas estrangeiras. Da mesma forma, profissionais que tem uma cultura básica de Estatística são muito valorizados exatamente pelo seu preparo para tomada de decisão. Mas o que significa isso?

→ Desenvolver uma cultura estatística significa desenvolver a habilidade de planejar um estudo controlando todos os aspectos que possam causar variações na resposta de interesse e, com base em metodologia científica ser capaz de realizar inferências que facilitem a tomada de decisão. Tem muito a ver com lógica e Bom Senso!

→ O profissional que trabalha com Estatística lida geralmente com uma amostra, ou seja, com uma pequena parcela da população, pela facilidade, pela limitação de tempo e custo mais baixo. Não precisamos comer a panela inteira de arroz para provar se está ou não bem temperada!

→ Apesar da Estatística ser considerada uma ciência exata, seus resultados estão sempre associados à uma pequena margem de erro, exatamente por estarem baseados em uma amostra. O profissional de estatística deve ter a habilidade de controlar esta margem de erro através de procedimentos de Amostragem.

→ A utilização de técnicas estatísticas foi incrementada com o grande avanço da computação eletrônica. Hoje existem centenas de softwares estatísticos que facilitam a utilização das técnicas. No entanto se não houver a habilidade de interpretação das informações do “output” do computador, a confusão estará formada!

→ Alguns exemplos de aplicação de técnicas estatísticas: Pesquisa Eleitoral, Pesquisa de Mercado, Controle de Qualidade, Índices Econômicos, além do desenvolvimento de medicamentos, formas de tratamento, sementes mais eficientes, etc, ou seja, tudo que é “comprovado cientificamente” deve passar por procedimentos estatísticos.

I – CONCEITOS E TIPOS DE PESQUISA

Definição:

A Pesquisa de Opinião é a investigação sistemática, controlada, empírica e crítica de dados com o objetivo de descobrir e/ou descrever fatos e/ou de verificar a existência de relações presumidas entre fatos (ou variáveis).

1.1 - Etapas de uma Pesquisa de Opinião

Étapas	Fases
1. Reconhecimento e formulação do problema de pesquisa	Formulação, determinação ou constatação de um problema de pesquisa;
2. Planejamento da pesquisa	a. Definição dos objetivos; b. Estabelecimento das questões de pesquisa; c. Estabelecimento das necessidades de dados e definição das variáveis e de seus indicadores; d. Determinação das fontes de dados; e. Determinação da metodologia; f. Planejamento da organização, cronograma e orçamento; g. Redação do projeto de pesquisa e/ou de proposta de pesquisa.
3. Execução da pesquisa	a. Preparação de campo; b. Campo; c. Processamento e análise.
4. Comunicação dos resultados	a. Elaboração e entrega dos relatórios de pesquisa; b. Preparação e apresentação oral dos resultados;

Reconhecimento e formulação do problema de pesquisa: consiste na correta identificação do problema de pesquisa que se pretenda resolver e que possa efetivamente receber contribuições valiosas da pesquisa de opinião em sua solução.

Planejamento da pesquisa: compreende a definição dos objetivos da pesquisa e de toda sua operacionalização. Fontes de dados, método de pesquisa, forma de coleta, construção e teste do instrumento de coleta, plano amostral, procedimentos de campo, plano de processamento e análise, definição dos recursos necessários, definição de cronograma das etapas.

Execução da pesquisa: coleta de dados e processamento, análise e interpretação.

Comunicação dos resultados: compreende a apresentação escrita e oral das principais descobertas da pesquisa, com sugestões e recomendações.

1.2- Tipos de Pesquisa

Projeto de Pesquisa: Cada planejamento de pesquisa realizado cientificamente tem um padrão específico para controlar a coleta de dados. Este padrão chama-se projeto de pesquisa. Sua função é assegurar que os dados exigidos sejam coletados de maneira precisa e econômica.

Os projetos de pesquisa podem ser agrupados nas seguintes categorias: exploratória, descritiva e experimental.

- **Pesquisa Exploratória:** visa fornecer ao pesquisador um maior conhecimento do tema ou problema de interesse. É apropriada para os primeiros estágios da investigação quando a familiaridade, o conhecimento e a compreensão do fenômeno por parte do pesquisador são insuficientes.

Alguns objetivos da Pesquisa Exploratória:

- Aumentar o conhecimento e compreensão de um problema de pesquisa;
- Auxiliar a formulação mais precisa do problema de pesquisa;
- Acumular informações disponíveis relacionadas ao problema de pesquisa;
- Ajudar na formulação de hipóteses a serem verificadas;
- Auxiliar a construção do instrumento de coleta de dados;

- Clarificar conceitos;
- Verificar pesquisas semelhantes já realizadas;
- Estabelecer prioridades para futuras pesquisas.

Métodos da Pesquisa Exploratória

O projeto formal está quase ausente nos estudos exploratórios. A imaginação do explorador é o fator principal. Entretanto, há 4 linhas de ataque que podem ajudar na descoberta de hipóteses valiosas:

Levantamentos em fontes secundárias: levantamentos bibliográficos, levantamentos documentais, levantamentos de estatísticas e levantamentos de pesquisas realizadas;

Levantamentos de experiências: Muitas pessoas, em função da posição estratégica que ocupam numa instituição acumulam experiências e conhecimentos sobre um tema ou problema em estudo. Informações são levantadas a partir de entrevistas individuais ou em grupo, realizadas com especialistas ou conhecedores do assunto.

Estudo de casos selecionados: exame de registros existentes, observação da ocorrência do fato, entrevistas, etc... (cases). Casos que reflitam mudanças, comportamentos ou desempenhos extremados, dificuldades superadas, etc.

Observação informal: A utilização do processo de observação do dia a dia em pesquisa exploratória deverá ser informal e dirigida, ou seja, centrada unicamente em observar objetos, comportamentos e fatos de interesse para o problema em estudo.

- **Pesquisa Descritiva:** destinam-se a descrever as características de determinada situação. A maioria das pesquisas de opinião realizadas são pesquisas descritivas. Ao contrário do que ocorre nas pesquisas exploratórias, a elaboração das questões de pesquisa pressupõe profundo conhecimento do problema a ser estudado. Um exemplo frequente é o estudo feito para definir o perfil de determinado público. São coletados dados para descobrir se as pessoas que possuem determinada característica estão concentradas no Norte ou no Sul, em cidades grandes ou pequenas, entre grupos de renda alta ou baixa, entre jovens ou pessoas idosas, etc.

Os estudos descritivos não devem ser encarados como simples coletas

de dados, embora infelizmente, muitos deles não sejam mais do que isso. Para ser valioso, o estudo descritivo precisa coletar dados com um objetivo definido e deve incluir uma interpretação por um investigador. Entre pessoas pouco familiarizadas com pesquisas de opinião, é comum construir rapidamente um questionário e sair a campo coletando milhares de dados. Em seguida, ficam decepcionadas com a inutilidade da maioria dos dados coletados dessa forma desordenada. A quantidade de opções de dados que podem ser coletados numa pesquisa é infinita, por isso é preciso saber antes, quais são os dados relevantes a serem coletados. Daí a necessidade do planejamento rigoroso da pesquisa.

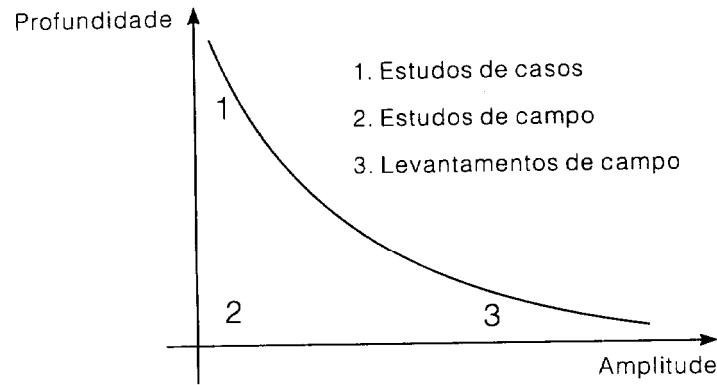
Alguns objetivos da pesquisa descritiva:

- Descrever as características de grupos de indivíduos;
- Estimar a proporção de elementos numa população específica que tenham determinadas características ou comportamentos;
- Descobrir ou verificar a existência de relação entre variáveis.

Tipos de pesquisa descritiva:

Levantamentos de campo (método estatístico) – grande amplitude e pouca profundidade – Procuram-se dados representativos da população de interesse, a amostra é gerada a partir de métodos estatísticos, tem-se total controle sobre a representatividade dos dados obtidos em relação à população. Permite a geração de tabelas sumarizadas por categorias e a generalização dos resultados para toda a população. No entanto não permite aprofundar os tópicos da pesquisa pela própria característica de gerar sumários estatísticos. É dispendioso em termos de tempo e isto e requer grandes conhecimentos técnicos.

Estudos de campo – média profundidade e média amplitude – É o método de estudo intensivo de um número relativamente pequeno de casos. Deve ser considerado como um estágio diferente no desenvolvimento de um método científico comum. Servem para geração de hipóteses em vez de teste de hipóteses, recomendados quando há grande homogeneidade entre os elementos da população. Entretanto somente investigam após a ocorrência do fato e geralmente não podem ser generalizados. Ex: Entrevista de profundidade.



Diferenciação entre estudos de casos, estudos de campo e levantamentos de campo, segundo a profundidade e a amplitude da pesquisa.

- **Pesquisa Experimental:** “Se ocorrer isto, provavelmente ocorrerá aquilo” Observação da relação de causalidade entre várias possíveis causas e o efeito pressuposto.

$$y = f(x, z, t, v, s, \dots)$$

onde y, variável dependente e as demais são independentes. Ganha-se maior confiabilidade nos resultados, à medida que repetidas experimentações com as mesmas variáveis independentes e dependente indicam sempre as mesmas conclusões.

1.3 - Fontes e tipos de dados

Há 4 tipos diferentes de fontes de dados em pesquisas de opinião:

- a) Pesquisado: os próprios indivíduos pesquisados são a maior fonte de dados em pesquisas de marketing, através de suas declarações, oralmente ou por escrito, ou através de sua observação;
- b) Pessoas que tenham informação sobre o pesquisado: quando ocorre do pesquisado ser inacessível é mais fácil conseguir a informação desejada com outras pessoas que convivem com ele.
- c) Situações similares: busca de conhecimento em situações análogas ou similares. As formas básicas de conhecer situações similares são através do estudo de casos, do experimento e da simulação;
- d) Dados disponíveis: existe uma infinidade de dados úteis e disponíveis para marketing que já foram coletados, catalogados e analisados. Podem ser encontrados em bibliotecas, órgãos governamentais, leitura de jornais e revistas, etc.

Tipos de dados:

Dados primários: são aqueles que não foram antes coletados e que são levantados com o propósito de atender às necessidades específicas da pesquisa em andamento. Podem ser características demográficas, socioeconômicas e de estilo de vida; atitudes e opiniões; conscientização e conhecimento; motivações; comportamento passado e presente; e intenções.

Dados secundários: são aqueles que já foram coletados, tabulados, ordenados e até analisados e que estão à disposição dos pesquisadores.

1.4- Critérios de Classificação Socioeconômica.

No Brasil, até 1970 não havia um critério único, objetivo e geral de classificação socioeconômica dos respondentes de pesquisa. A partir daí esforços foram despendidos no sentido de se construir critérios confiáveis e facilmente implementáveis. Em vigor, atualmente se utiliza o Critério Brasil 2015 (ABEP):

Variáveis	Quantidade				
	0	1	2	3	4 ou +
Banheiros	0	3	7	10	14
Empregados domésticos	0	3	7	10	13
Automóveis	0	3	5	8	11
Microcomputador	0	3	6	8	11
Lava louca	0	3	6	6	6
Geladeira	0	2	3	5	5
Freezer	0	2	4	6	6
Lava roupa	0	2	4	6	6
DVD	0	1	3	4	6
Micro-ondas	0	2	4	4	4
Motocicleta	0	1	3	3	3
Secadora roupa	0	2	2	2	2

Grau de instrução do chefe de família e acesso a serviços públicos Escolaridade da pessoa de referência

Grau de instrução	Pontos	
Analfabeto / Fundamental I incompleto	0	
Fundamental I completo / Fundamental II incompleto	1	
Fundamental II completo / Médio incompleto	2	
Médio completo / Superior incompleto	4	
Superior completo	7	
Serviços públicos		
	Não	Sim
Água encanada	0	4
Rua pavimentada	0	2

Cortes do Critério Brasil Classe	Pontos
A	45 - 100
B1	38 - 44
B2	29 - 37
C1	23 - 28
C2	17 - 22
D-E	0 - 16

II – MEDIDAS E INSTRUMENTOS DE COLETA DE DADOS

2.1 - Método científico aplicado à pesquisa mercadológica

“Nenhum método conhecido pelo homem pode eliminar inteiramente a incerteza. Porém, o método científico, mais do que qualquer outro processo, pode diminuir os elementos de incerteza que resultam da falta de informação. Assim fazendo, reduz o perigo de uma escolha errada entre várias alternativas” (JAHODA, DEUTSCH & COOK, 1951)

Pesquisa de Opinião é a aplicação do método científico na caracterização de um perfil populacional. O método científico coincide com a definição de pesquisa de opinião: coleta, registro e análise de fatos sobre questões de interesse. Entretanto, além disso, o método científico salienta que isso precisa ser feito com cuidado e precisão. Existem 3 características que diferenciam o método científico do não científico: a objetividade do investigador, a precisão da medida e a natureza contínua e exaustiva da investigação.

O ser humano, como objeto de estudo, é extremamente complexo, especialmente a sua capacidade de raciocínio e de agir conforme a sua própria vontade. As informações coletadas numa pesquisa são geralmente obtidas por entrevistas, um processo subjetivo que não permite medidas precisas; são opiniões e atitudes que somente podem ser quantificadas por meio de aproximações bastante amplas. Desta forma ainda há uma dúvida se a pesquisa de opinião pode ser chamada científica, pela dificuldade de se medir a precisão. Deve-se obter as medidas mais precisas possíveis com os métodos disponíveis no campo de estudo específico.

2.2 - Tipos básicos de escalas

Escala nominal: são aquelas em que os números são utilizados apenas para nomear, identificar e/ou categorizar dados sobre pessoas, objetos ou fatos. São intensivamente utilizadas em pesquisas de opinião, em variáveis como: marcas, cores, modelos, sexo, tipo de moradia, regiões, uso/não uso, gosta/não gosta, ocupação, etc. Não faz nenhum sentido calcular medidas descritivas em escalas nominais. Deve-se tomar cuidado ao se trabalhar com o computador, pois este não identifica a escala de mensuração das variáveis.

Escala ordinal: é aquela em que os números servem para, além de nomear, identificar e/ou categorizar, ordenar segundo um processo de comparação, as pessoas, objetos ou fatos, em relação a determinada característica. Esta

escala nos permite afirmar que o produto A é melhor que o produto B, entretanto não nos permite saber o quanto A é melhor que B.

Escala intervalar: é aquela que ocorre quando a escala tem as características da escala ordinal e ainda é possível quantificar a diferença entre dois números desta escala. Entretanto não nos permite concluir quanto à magnitude absoluta das medições. Em pesquisas de opinião, são muito usadas para medir atitudes, opiniões, conscientização e preferências, onde a preocupação é estabelecer medidas relativas e não absolutas. Ex: temperatura

2.3 - Instrumentos de coleta de dados

O instrumento de coleta de dados é o documento através do qual as perguntas e questões serão apresentadas aos respondentes e onde são registradas as respostas e dados obtidos. O questionário e a observação são os dois métodos básicos de coleta de dados em pesquisas de opinião. No questionário, os dados são coletados por meio de perguntas a pessoas que se supõe terem a informação desejada. As questões podem ser feitas pessoalmente ou por escrito. Quando os dados são coletados por observação, o pesquisador não faz perguntas. Ele observa as coisas ou atos nos quais está interessado.

Questionário: O método do questionário se tornou o mais largamente usado dos dois métodos de coletar dados. Hoje as pessoas estão mais familiarizadas com abordagem de entrevistadores a qualquer hora do dia. É um instrumento versátil, quase todos os problemas de pesquisa podem ser abordados através de um questionário. É normalmente mais rápido e mais barato do que a observação. Entretanto, este método ainda esbarra na má vontade do entrevistado para dar informações, ou na incapacidade de fornecer informações precisas, seja por esquecimento ou por falta de informação. Muitas vezes usa-se o artifício de se estimular a resposta do entrevistado, facilitando assim a obtenção das informações. E em alguns casos, o entrevistado tende a querer “agradar” o pesquisador declarando uma resposta que imagina seja de interesse do pesquisador. Sem falar nas interpretações errôneas que muitas vezes ocorrem na leitura das questões causadas pelas dificuldades semânticas. Três aspectos são bastante importantes no processo de construção de um questionário: 1. o grau em que o questionário é formalizado ou estruturado, 2. o disfarce ou falta de disfarce dos objetivos do questionário e 3. o método de comunicação usado.

Estrutura	Estruturado	Não-estruturado
Disfarce		
Disfarçado	O conhecimento do indivíduo, sua percepção e memória estão condicionados por suas atitudes.	O entrevistado não sabe o motivo do estudo. São usadas as técnicas de associação de idéias, completar sentenças e contar histórias.
Não-disfarçado	Maioria dos questionários. As respostas são frequentemente limitadas a uma lista de alternativas que estão implícitas ou declaradas.	O entrevistador tenta fazer com que o entrevistado fale livremente a respeito do assunto de interesse. Entrevista de profundidade.

Existem 3 métodos diferentes de comunicação por questionários: 1. entrevista pessoal, 2. por telefone, 3. pelo correio/email.

Um exercício interessante é promover uma discussão a respeito das vantagens dos 3 métodos segundo os fatores: Flexibilidade, Quantidade de informação, Precisão da informação obtida, Amostragem, Velocidade, Custo e Administração.

As perguntas dos questionários podem ser: abertas ou fechadas (dicotômicas, escolha única, escolha múltipla ou escala).

Observação: Teoricamente esse método é superior ao método do questionário. Os acontecimentos são registrados à medida que ocorrem e a influência do entrevistador e do questionário na resposta do entrevistado é eliminada, tornando-a mais precisa e objetiva. Alguns estudos apontam grandes diferenças nos resultados obtidos através dos dois métodos para uma mesma pesquisa.

Por outro lado, a falta de versatilidade do método da observação já que não se podem obter dados diretos sobre as atitudes dos indivíduos. O custo também é um fator complicador, pois é necessário colocar observadores onde possam observar o fenômeno, isto requer uma espera do observador e este tempo improdutivo é um custo adicional. Portanto, do ponto de vista prático, as vantagens são a favor do método do questionário.

Muitos estudos usam os dois métodos de coleta de dados. Enquanto fazem perguntas, os entrevistadores frequentemente observam a idade, sexo, indicações de nível sócio econômico e outros dados.

III – AMOSTRAGEM E INTERVALO DE CONFIANÇA

3.1- Conceito de Amostragem

AMOSTRA ALEATÓRIA

Considera-se uma população e observando alguns elementos queremos obter informações a respeito da mesma. Estes elementos são obtidos de forma aleatória (em geral independentes) o que é chamada de *amostra aleatória*.

É preciso garantir que a amostra ou as amostras que serão usadas sejam obtidas por processos adequados. Se erros forem cometidos no momento de selecionar os elementos da amostra, o trabalho todo ficará comprometido e os resultados finais serão provavelmente bastante viesados. Devemos, portanto, tomar especial cuidado quanto aos critérios que serão usados na seleção da amostra.

O que é necessário garantir, em suma, é que a amostra seja representativa da população. Isso significa que, com exceção de pequenas discrepâncias inerentes à 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) que desejamos pesquisar.

A necessidade da representatividade da amostra não é difícil de entender. O que talvez não seja fácil é saber quando temos uma amostra representativa ou não. Veremos adiante algumas recomendações sobre como proceder para garantir, da melhor forma possível, a representatividade da amostra.

Os problemas de amostragem podem ser mais ou menos complexos, dependendo das populações e das variáveis que se deseja estudar. Na indústria, onde amostras são frequentemente retiradas para efeito de controle de qualidade dos produtos e materiais, em geral os problemas de amostragem são mais simples de resolver. Por outro lado, em pesquisas sociais, econômicas ou de opinião, a complexidade dos problemas de amostragem é normalmente bastante grande. Em tais casos, extremo cuidado deve ser tomado quanto à caracterização da população e ao processo usado para selecionar a amostra, a fim de evitar que os elementos desta constituam um conjunto com características fundamentalmente distintas das da população.

No caso de questionários serem distribuídos, muita atenção é também requerida em sua elaboração, visando evitar perguntas capciosas ou inibidoras, o que viria a distorcer os resultados.

Em resumo, a obtenção de soluções adequadas para o problema de amostragem exige, em geral, muito bom senso e experiência. Além disso, é muitas vezes conveniente que o trabalho seja acompanhado por um especialista do assunto em questão.

Veremos a seguir, algumas recomendações básicas referentes ao problema de amostragem e a apresentação das principais técnicas de amostragem.

3.2 - Técnicas de amostragem

AMOSTRAGEM PROBABILÍSTICA

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

A amostragem será probabilística se todos os elementos da população tiverem probabilidade conhecida, e diferente de zero, de pertencer à amostra. Caso contrário, a amostragem será não probabilística.

Segundo essa definição, a amostragem probabilística implica um sorteio com regras bem determinadas, cuja realização só será possível se a população for finita e totalmente acessível.

A utilização de uma amostragem probabilística é a melhor recomendação que se deve fazer no sentido de se garantir a representatividade da amostra, pois o acaso será o único responsável por eventuais discrepâncias entre população e amostra.

A seguir, damos algumas das principais técnicas de amostragem probabilística.

AMOSTRAGEM ALEATÓRIA SIMPLES

Esse tipo de amostragem, também chamada *simples ao acaso, casual, elementar, randômica, etc.*, é equivalente a um sorteio lotérico. Nela, todos os elementos da população têm igual probabilidade de pertencer à amostra, e todas as possíveis amostras têm também igual probabilidade de ocorrer.

Sendo N o número de elementos da população e n o número de elementos da amostra, cada elemento da população tem probabilidade $\frac{n}{N}$ de pertencer à amostra. A essa relação $\frac{n}{N}$ denomina-se *fração de amostragem*. Por outro lado, sendo a amostragem feita sem reposição, o que suporemos em geral, existem $\binom{N}{n}$ possíveis amostras, todas igualmente prováveis.

Na prática, a amostragem simples ao acaso pode ser realizada numerando-se a população de 1 a N , sorteando-se, a seguir, por meio de um dispositivo aleatório qualquer, n números dessa sequência, os quais corresponderão aos elementos sorteados para a amostra. Esse sorteio pode ser feito através da geração *de números aleatórios*.

AMOSTRAGEM SISTEMÁTICA

Quando os elementos da população se apresentam ordenados e a retirada dos elementos da amostra é feita periodicamente, temos uma *amostragem sistemática*.

Assim, por exemplo, em uma linha de produção, podemos, a cada dez itens produzidos, retirar um para pertencer a uma amostra da produção diária.

A principal vantagem da amostragem sistemática está na grande facilidade na determinação dos elementos da amostra. O perigo em adotá-la está na possibilidade da existência de ciclos de variação da variável de interesse, 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 casual simples, podendo ser utilizada sem restrições.

AMOSTRAGEM ESTRATIFICADA

Muitas vezes a população se divide em subpopulações ou estratos, sendo razoável supor que, de estrato para estrato, a variável de interesse apresente um comportamento substancialmente diverso, tendo, entretanto, comportamento razoavelmente homogêneo dentro de cada estrato. Em tais casos, se o sorteio dos elementos da amostra for realizado sem se levar em consideração a existência dos estratos, pode acontecer que os diversos estratos não sejam convenientemente representados na amostra, a qual seria mais influenciada pelas características da variável nos estratos mais favorecidos pelo sorteio. Evidentemente, a tendência à ocorrência de tal fato será tanto maior quanto menor o tamanho da amostra. Para evitar isso, pode-se adotar uma *amostragem estratificada*.

Constituem exemplos em que uma amostragem estratificada parece ser recomendável, a estratificação de uma cidade em bairros, quando se deseja investigar alguma variável relacionada à renda familiar; a estratificação de uma população humana em homens e mulheres, ou por faixas etárias; a estratificação de uma população de estudantes conforme suas especificações, etc.

AMOSTRA DE CONVENIÊNCIA

A *amostra de conveniência* é formada por elementos que o pesquisador reuniu simplesmente porque dispunha deles. Então, se o professor tomar os alunos de sua classe como amostra de toda a escola, estará usando uma amostra de conveniência.

Os estatísticos têm muitas restrições ao uso de amostras de conveniência. Mesmo assim, as amostras de conveniência são comuns na área de saúde, onde se fazem pesquisas com pacientes de uma só clínica ou de um só hospital. Mais ainda, as amostras de conveniência constituem, muitas vezes, a única maneira de estudar determinado problema.

De qualquer forma, o pesquisador que utiliza amostras de conveniência precisa de muito senso crítico. Os dados podem ser tendenciosos. Por exemplo, para estimar a probabilidade de morte por desidratação não se deve recorrer aos dados de um hospital. Como só são internados os casos graves, é possível que a mortalidade entre pacientes internados seja maior do que entre pacientes não-internados. Consequentemente, a amostra de conveniência, constituída, neste exemplo, por pacientes internados no hospital, seria tendenciosa.

Finalmente, o pesquisador que trabalha com amostras sempre pretende fazer inferência, isto é, estender os resultados da amostra para toda a população. Então é muito importante caracterizar bem a amostra e estender os resultados obtidos na amostra apenas para a população de onde a amostra proveio.

3.3 - Cálculo do tamanho amostral;

ERRO AMOSTRAL E NÍVEL DE CONFIANÇA

Uma das etapas do planejamento amostral é a definição do erro amostral e do nível de confiança desejado pelo pesquisador, no levantamento a ser realizado. Estas são informações que poderão levar à um aumento ou redução do tamanho amostral, dependendo dos recursos disponíveis para realização da pesquisa. Trataremos abaixo desses conceitos especificamente para o caso em que estamos interessados em estimar uma proporção populacional.

1) ERRO AMOSTRAL

Imagine que para estimação de uma proporção (ou porcentagem) tenhamos interesse em atingir uma certa precisão para nossa estimativa de mais ou menos ϵ %. Isto significa que se o nosso levantamento proporcionar uma estimativa de P % então poderemos afirmar que a porcentagem da população deve estar entre

$$P - \epsilon < P < P + \epsilon \quad (1)$$

com uma certa margem (ou nível) de confiança;

2) NÍVEL DE CONFIANÇA

Imagine agora que o levantamento amostral acima tenha sido realizado com a definição de um nível de confiança de $(1 - \alpha)$ ou $(1 - \alpha) \times 100$ %. Isto significa que se realizarmos 100 levantamentos semelhantes, o intervalo definido acima (1) conterá o verdadeiro valor da população em $(1 - \alpha) \times 100$ % das vezes.

O nível de confiança é exatamente o oposto do nível de significância (α) de forma que o efeito causado pelo aumento ou diminuição dessas quantidades pode ser melhor representado na tabela abaixo:

Erro amostral	Nível de confiança	Nível de significância	Tamanho amostral
↓	↑	↓	↑
↑	↓	↑	↓

CÁLCULO DE TAMANHO AMOSTRAL

Desta forma uma expressão que poderá nos ajudar na definição do tamanho que uma amostra deve ter para satisfazer algumas condições pré-estabelecidas é dada por:

$$n_0 = \frac{Z^2 pq}{\varepsilon^2}$$

onde:

Z é o valor da variável padronizada para um nível de significância α . É o valor da distribuição Normal padronizada cuja área à direita é de α %;

p é uma pré-estimativa definida a partir de informações que o pesquisador possui acerca desse parâmetro. Geralmente é desconhecida e, portanto, assume-se que ela é igual a 0,5;

q é o complemento de p, ou seja, $q = 1 - p$;

ε é o erro amostral definido pelo pesquisador;

n_0 é o tamanho amostral sugerido.

Sugere-se ainda que seja realizada uma correção no caso de estarmos trabalhando com uma população finita, da seguinte forma:

$$n = \frac{n_0}{1 + \left(\frac{n_0}{N} \right)}$$