

Ministério da Educação – MEC
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES
Diretoria de Educação a Distância – DED
Universidade Aberta do Brasil – UAB
Programa Nacional de Formação em Administração Pública – PNAP
Bacharelado em Administração Pública

ESTATÍSTICA APLICADA À ADMINISTRAÇÃO

Marcelo Tavares



2011

© 2011. Universidade Federal de Santa Catarina – UFSC. Todos os direitos reservados.

A responsabilidade pelo conteúdo e imagens desta obra é do(s) respectivo(s) autor(es). O conteúdo desta obra foi licenciado temporária e gratuitamente para utilização no âmbito do Sistema Universidade Aberta do Brasil, através da UFSC. O leitor se compromete a utilizar o conteúdo desta obra para aprendizado pessoal, sendo que a reprodução e distribuição ficarão limitadas ao âmbito interno dos cursos. A citação desta obra em trabalhos acadêmicos e/ou profissionais poderá ser feita com indicação da fonte. A cópia desta obra sem autorização expressa ou com intuito de lucro constitui crime contra a propriedade intelectual, com sanções previstas no Código Penal, artigo 184, Parágrafos 1º ao 3º, sem prejuízo das sanções cíveis cabíveis à espécie.

T231e	Tavares, Marcelo Estatística aplicada à administração / Marcelo Tavares. – Florianópolis : Departamento de Ciências da Administração / UFSC; [Brasília] : CAPES : UAB, 2011. 222p. : il. Inclui bibliografia Bacharelado em Administração Pública ISBN: 978-85-7988-099-5 1. Administração – Métodos estatísticos. 2. Estatística. 3. Probabilidades. 4. Educação a distância. I. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Brasil). II. Universidade Aberta do Brasil. III. Título. CDU: 519.2:65
-------	--

Catálogo na publicação por: Onélia Silva Guimarães CRB-14/071

PRESIDENTA DA REPÚBLICA

Dilma Vana Rousseff

MINISTRO DA EDUCAÇÃO

Fernando Haddad

PRESIDENTE DA CAPES

Jorge Almeida Guimarães

UNIVERSIDADE FEDERAL DE SANTA CATARINA

REITOR

Alvaro Toubes Prata

VICE-REITOR

Carlos Alberto Justo da Silva

CENTRO SÓCIO-ECONÔMICO

DIRETOR

Ricardo José de Araújo Oliveira

VICE-DIRETOR

Alexandre Marino Costa

DEPARTAMENTO DE CIÊNCIAS DA ADMINISTRAÇÃO

CHEFE DO DEPARTAMENTO

Gilberto de Oliveira Moritz

SUBCHEFE DO DEPARTAMENTO

Marcos Baptista Lopez Dalmau

DIRETORIA DE EDUCAÇÃO A DISTÂNCIA

DIRETORIA DE EDUCAÇÃO A DISTÂNCIA

COORDENAÇÃO GERAL DE ARTICULAÇÃO ACADÊMICA

Liliane Carneiro dos Santos Ferreira

COORDENAÇÃO GERAL DE SUPERVISÃO E FOMENTO

Grace Tavares Vieira

COORDENAÇÃO GERAL DE INFRAESTRUTURA DE POLOS

Joselino Goulart Junior

COORDENAÇÃO GERAL DE POLÍTICAS DE INFORMAÇÃO

Adi Balbinot Junior

COMISSÃO DE AVALIAÇÃO E ACOMPANHAMENTO – PNAP

Alexandre Marino Costa

Claudinê Jordão de Carvalho

Eliane Moreira Sá de Souza
Marcos Tanure Sanabio
Maria Aparecida da Silva
Marina Isabel de Almeida
Oreste Preti
Tatiane Michelin
Teresa Cristina Janes Carneiro

METODOLOGIA PARA EDUCAÇÃO A DISTÂNCIA

Universidade Federal de Mato Grosso

COORDENAÇÃO TÉCNICA – DED

Tatiane Michelin
Tatiane Pacanaro Trinca
Soraya Matos de Vasconcelos

AUTOR DO CONTEÚDO

Marcelo Tavares

EQUIPE DE DESENVOLVIMENTO DE RECURSOS DIDÁTICOS CAD/UFSC

Coordenador do Projeto

Alexandre Marino Costa

Coordenação de Produção de Recursos Didáticos

Denise Aparecida Bunn

Supervisão de Produção de Recursos Didáticos

Érika Alessandra Salmeron Silva

Designer Instrucional

Adriana Novelli

Andreza Regina Lopes da Silva

Denise Aparecida Bunn

Auxiliar Administrativo

Stephany Kaori Yoshida

Capa

Alexandre Noronha

Projeto Gráfico e Finalização

Annye Cristiny Tessaro

Diagramação

Rita Castelan Minatto

Ilustração

Livia Remor Pereira

Revisão Textual

Claudia Leal Estevão Brites Ramos

Mara Aparecida Andrade R. Siqueira

PREFÁCIO

Os dois principais desafios da atualidade na área educacional do País são a qualificação dos professores que atuam nas escolas de educação básica e a qualificação do quadro funcional atuante na gestão do Estado brasileiro, nas várias instâncias administrativas. O Ministério da Educação (MEC) está enfrentando o primeiro desafio com o Plano Nacional de Formação de Professores, que tem como objetivo qualificar mais de 300.000 professores em exercício nas escolas de Ensino Fundamental e Médio, sendo metade desse esforço realizado pelo Sistema Universidade Aberta do Brasil (UAB). Em relação ao segundo desafio, o MEC, por meio da UAB/CAPES, lança o Programa Nacional de Formação em Administração Pública (PNAP). Esse Programa engloba um curso de bacharelado e três especializações (Gestão Pública, Gestão Pública Municipal e Gestão em Saúde) e visa a colaborar com o esforço de qualificação dos gestores públicos brasileiros, com especial atenção no atendimento ao interior do País, por meio dos Polos da UAB.

O PNAP é um Programa com características especiais. Em primeiro lugar, surgiu do esforço e da reflexão de uma rede composta pela Escola Nacional de Administração Pública (ENAP), pelo Ministério do Planejamento, pelo Ministério da Saúde, pelo Conselho Federal de Administração, pela Secretaria de Educação a Distância (SEED) e por mais de 20 instituições públicas de Ensino Superior (IPES), vinculadas à UAB, que colaboraram na elaboração do Projeto Político Pedagógico dos cursos. Em segundo lugar, esse Projeto será aplicado por todas as instituições e pretende manter um padrão de qualidade em todo o País, mas abrindo margem para

que cada IPES, que ofertará os cursos, possa incluir assuntos em atendimento às diversidades econômicas e culturais de sua região.

Outro elemento importante é a construção coletiva do material didático. A UAB colocará à disposição das IPES um material didático mínimo de referência para todas as disciplinas obrigatórias e para algumas optativas. Esse material está sendo elaborado por profissionais experientes da área da Administração Pública de mais de 30 diferentes instituições, com o apoio de equipe multidisciplinar. Por último, a produção coletiva antecipada dos materiais didáticos libera o corpo docente IPES para uma dedicação maior ao processo de gestão acadêmica dos cursos; uniformiza um elevado patamar de qualidade para o material didático e garante o desenvolvimento ininterrupto dos cursos, sem paralisações que sempre comprometem o entusiasmo dos estudantes.

Por tudo isso, estamos seguros de que mais um importante passo em direção à democratização do Ensino Superior público e de qualidade está sendo dado, desta vez contribuindo também para a melhoria da gestão pública brasileira.

Celso José da Costa
Diretor de Educação a Distância
Coordenador Nacional da UAB
CAPES-MEC

SUMÁRIO

Apresentação.....	11
-------------------	----

Unidade 1 – Fases do Método Estatístico, População e Amostra

Fases do Método Estatístico.....	17
Definição do Problema.....	18
Planejamento.....	19
Coleta de Dados.....	19
Organização de Dados.....	20
Apresentação de Dados.....	21
Análise e Interpretação de Dados.....	22
População e Amostras.....	25
Amostragens Probabilísticas.....	31
Amostragens não Probabilísticas.....	40

Unidade 2 – Distribuições de Frequências e Representação Gráfica

Distribuições de Frequências.....	49
Distribuições de Frequências de uma Variável Quantitativa Contínua.....	51
Distribuições de Frequências de uma Variável Qualitativa.....	58
Distribuições de Frequências de uma Variável Quantitativa Discreta.....	58
Representação Gráfica.....	61

Unidade 3 – Medidas de Posição e Dispersão

Medidas de Posição.....	73
Média.....	74
Moda.....	78
Mediana.....	80
Separatrizes.....	82
Medidas de Dispersão.....	85
Amplitude Total.....	86
Variância.....	87
Desvio Padrão.....	89
Coeficiente de Variação.....	90

Unidade 4 – Probabilidade

Introdução.....	101
Experimento Aleatório.....	103
Espaço Amostral.....	105
Evento.....	106
Definições de Probabilidades.....	107
Probabilidade Condicional.....	113
Regra do Produto e Eventos Independentes.....	116
Algumas Regras Básicas de Probabilidades.....	122

Unidade 5 – Distribuição de Probabilidades Discretas e Contínuas

Introdução.....	129
Distribuições Discretas.....	131
Distribuição Binomial.....	132
Distribuição de Poisson.....	136

Distribuições Contínuas.....	140
Distribuição Normal.....	140
Distribuições Amostrais.....	148
Distribuição t de <i>Student</i>	149
Distribuição de Qui-Quadrado.....	153
Distribuição de F.....	156
Noções de Estimação.....	159
Estimação por Intervalos.....	161
Dimensionamento de Amostras.....	164

Unidade 6 – Testes de Hipóteses

Introdução.....	171
Estrutura dos Testes de Hipóteses.....	173
Teste de Hipótese para uma Média.....	178
Teste de Hipótese para a Razão de duas Variâncias.....	183
Teste de Hipótese para a Diferença entre Médias.....	185
Teste de Hipótese para a Diferença entre Proporções.....	198
Teste de Qui-Quadrado de Independência.....	201
Associação entre Variáveis.....	205

Considerações finais	219
-----------------------------------	-----

Referências	220
--------------------------	-----

Minicurriculo	222
----------------------------	-----

APRESENTAÇÃO

Seja bem-vindo ao estudo da Estatística!

Esperamos que esta disciplina seja uma experiência interessante e enriquecedora. Pensando nisso, elaboramos o material com cuidado para que você aprenda os principais conceitos associados à Estatística, que vem se tornando cada vez mais importante no competitivo ambiente de negócios e de gestão. Juntos, iremos viajar pelo mundo dos números associados à estatística e suas relações no dia a dia do gestor público.

O principal objetivo é que você tenha a oportunidade de ampliar seu conhecimento sobre o universo da estatística. Dessa forma, não serão feitas neste material deduções e demonstrações matemáticas de expressões. Será tentado um entendimento mais abstrato das expressões a serem utilizadas.

Você já deve estar acostumado a utilizar a estatística, ou ferramentas estatísticas, no seu dia a dia, sem saber que a está utilizando. Se você acha que a estatística se resume apenas a números e a gráficos, está redondamente enganado. Dessa forma, estaremos, a partir de agora, entrando em um mundo no qual os números irão sempre lhe falar ou lhe contar alguma coisa. O seu trabalho usando a estatística passará a ser o de ajudar a planejar a obtenção de dados, a interpretar e a analisar os dados obtidos e a apresentar os resultados de maneira a facilitar a sua tomada de decisões como gestor na área pública.

Para gerar tabelas, gráficos e utilizar técnicas estatísticas, temos uma infinidade de *softwares* que fazem isso automaticamente. Entretanto, para podermos descobrir quais as respostas que os dados podem nos dar para determinados questionamentos, é necessário

que você saiba a teoria estatística e treine suas aplicações por meio de estudos de casos, ou situações.

Sempre surgem, então, perguntas do tipo: quais variáveis devem ser medidas? Como retirar amostras da população que se deseja estudar? Que tipo de análise realizar? Como interpretar os resultados? Espero que no final da leitura deste material você tenha condições de responder de forma clara a essas perguntas e a outras que podem ser feitas.

É necessário termos em mente que a estatística é uma ferramenta para o gestor ou para o executivo, nas respostas aos “porquês” de seus problemas. Contudo, para que ela seja bem utilizada, é necessário conhecer os seus fundamentos e os seus princípios e, acima de tudo, que o gestor desenvolva um espírito crítico e de análise; pois é fácil mentir usando a estatística, difícil é falar a verdade sem usar a estatística.

Atualmente, as empresas têm procurado admitir como gestores os profissionais que tenham um nível de conhecimento de estatística alto, pois esse conhecimento tem resultado em diferença significativa nos processos decisórios.

Para estudar na modalidade a distância o conteúdo da disciplina *Estatística Aplicada à Administração* é preciso que você tenha disciplina intelectual que, para desenvolver, somente praticando e, ainda, uma postura crítica, sistemática. Ou seja, ao invés de você atuar como passivo e concordar com tudo o que diz o texto, você deve **duvidar, contestar, criticar, comentar e descobrir o que o autor quer dizer**. O ato de estudar exige que você **faça exercícios e entenda o que está fazendo**, não sendo apenas um mero executor de fórmulas. Isso implica o entendimento dos conceitos apresentados neste material.

Uma vez que a leitura é uma atividade, você deve ser ativo. Tenha certeza de que um estudante consegue aprender mais do que o outro à medida que se aplica mais e é capaz de uma atividade maior de leitura. E aprende melhor se exigir mais de si mesmo e do texto que tem diante de si.

Para tanto, dividimos o livro em seis Unidades. Na Unidade 1, iremos ver as fases do método estatístico e os conceitos de populações, de amostras e de métodos de amostragem. Nas Unidades 2 e 3, vamos aprender a descrever um conjunto de dados por meio de distribuições de frequências, de medidas de posição e de dispersão. Já nas Unidades 4 e 5, estaremos aprendendo conceitos relacionados a probabilidades, a distribuições discretas e contínuas, além de noções de estimação. E, por fim, na última Unidade, você irá aprender como tomar decisões baseadas nos chamados testes de hipóteses.

Desejo a você bons estudos!

Professor Marcelo Tavares

UNIDADE 1

FASES DO MÉTODO ESTATÍSTICO, POPULAÇÃO E AMOSTRA

OBJETIVOS ESPECÍFICOS DE APRENDIZAGEM

Ao finalizar esta Unidade, você deverá ser capaz de:

- ▶ Entender as relações entre as fases do método estatístico e aplicá-las no desenvolvimento de seus projetos;
- ▶ Compreender conceitos básicos relacionados à estatística, como variáveis, estimadores, estimativas, parâmetros, população, amostras; e
- ▶ Entender os diversos tipos de amostragem e saber como aplicá-los quando for desenvolver qualquer tipo de projeto em que se utilize planos amostrais.

FASES DO MÉTODO ESTATÍSTICO

Caro estudante,

Vamos iniciar nossos estudos de estatística para que você tenha condições de identificar a forma pela qual podemos utilizá-la, seja dentro da pesquisa científica ou na tomada de decisões e na estruturação de projetos.

Além disso, trabalharemos a definição de população e de amostra, bem como a forma de retirar as amostras de uma população; temas de fundamental importância para que você consiga desenvolver trabalhos com resultados de campo de alto nível.

Vamos então aprender esses assuntos? Boa leitura e, qualquer dúvida, não hesite em consultar o seu tutor.

Para realizarmos um estudo estatístico, normalmente, existem várias etapas a serem realizadas. Essas etapas são chamadas de fases do método estatístico. Quando você tiver bem definido essas fases, e tiver condições de realizá-las de forma adequada, a chance de sucesso em um trabalho estatístico ou que envolva estatística será muito maior. Para isso, então, você irá conhecer essas fases ou etapas de forma mais detalhada.

As fases do método estatístico são:

- ▶ definição do problema;
- ▶ planejamento do processo de resolução;
- ▶ coleta dos dados;
- ▶ organização de dados;
- ▶ apresentação de dados; e
- ▶ análise e interpretação dos resultados.

Agora, você verá de forma detalhada cada uma dessas fases. Ao longo da apresentação, iremos detalhando cada uma delas, inserindo-as passo a passo, para que no final você tenha uma ideia das relações entre essas fases.

DEFINIÇÃO DO PROBLEMA

A primeira fase consiste na definição e na formulação correta do problema a ser estudado. Para isso, você deve procurar outros estudos realizados sobre o tema escolhido, pois, desse modo, você evitará cometer erros que já tenham sido cometidos por outros.

Sendo assim, essa primeira fase pode responder à definição de um problema ou, simplesmente, dar resposta a um interesse de profissionais. Em alguns casos, podem estar envolvidas **variáveis qualitativas e quantitativas**, por exemplo:

Veremos esses conceitos mais adiante nesta Unidade, ao longo da disciplina.

- ▶ a receita do Imposto Territorial e Predial Urbano (IPTU) de cada um dos bairros de uma cidade em vários anos;
- ▶ o tratamento de dados relativos aos desempenhos dos funcionários de um setor de uma prefeitura ao longo de alguns meses;
- ▶ a quantidade de residências em uma cidade que atrasam o pagamento do IPTU em 1, 2, 3, 4, 5 ou mais meses; e
- ▶ o tempo necessário entre o pedido de reparo de uma via pública e a realização do serviço.

Mas não para por aí! Existem outros problemas relacionados à gestão pública que merecem ser resolvidos.

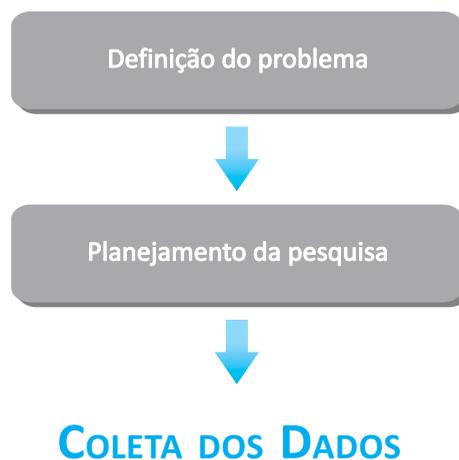
Definição do problema



PLANEJAMENTO

Após você definir o problema, é preciso determinar um processo para resolvê-lo e, em especial, a forma de como obter informações sobre a variável ou as variáveis em estudo. É nessa fase que devemos decidir pela observação da população ou de uma amostra. Portanto, você precisa:

- ▶ determinar os procedimentos necessários para resolver o problema, em especial, como levantar informações sobre o assunto objeto do estudo;
- ▶ planejar o trabalho tendo em vista o objetivo a ser atingido;
- ▶ escolher e formular corretamente as perguntas;
- ▶ definir o tipo de levantamento – censitário ou por amostragem; e
- ▶ definir cronograma de atividades, custos envolvidos, delineamento da amostra etc.



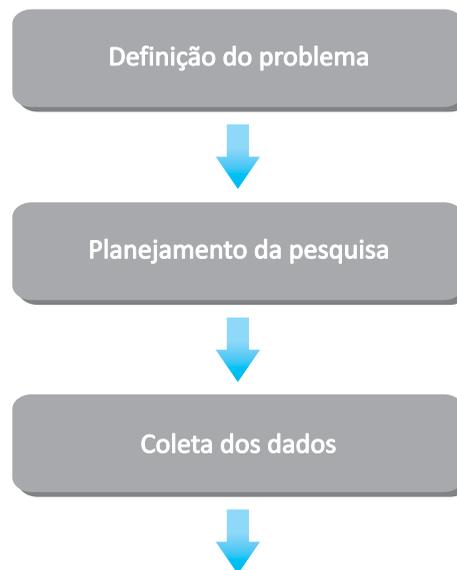
Agora que você já planejou o seu trabalho, vamos para a terceira etapa, que consiste na coleta de dados. Essa fase que deve ser seguida com cuidado, pois dados mal coletados resultam em

estatísticas inadequadas ou que não refletem a situação que você deseja estudar.

Os dados podem ser coletados, por exemplo, por meio de:

- ▶ questionário;
- ▶ observação;
- ▶ experimentação; e
- ▶ pesquisa bibliográfica.

A coleta de dados que você vai fazer pode ser realizada de forma direta com base nos elementos de registros ou pelo próprio pesquisador através de questionários.

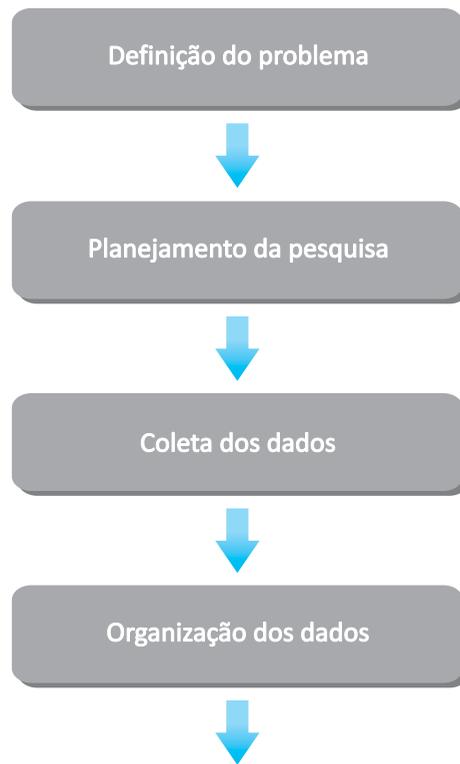


ORGANIZAÇÃO DE DADOS

***Atributo estatístico** – é toda medida estatística, por exemplo, média. Fonte: Elaborado pelo autor.

***Tratamento estatístico** – implica analisar os dados utilizando técnicas estatísticas. Fonte: Elaborado pelo autor.

Agora que temos os dados precisamos organizá-los, pois somente coletar os dados não é suficiente. Essa organização consiste em “resumir” os dados através da sua contagem e agrupamento. Desse modo, obtemos um conjunto de informações que irá conduzir ao estudo do **atributo estatístico***. Geralmente, essa organização é feita em planilhas eletrônicas (tipo Excel) para posterior **tratamento estatístico***.

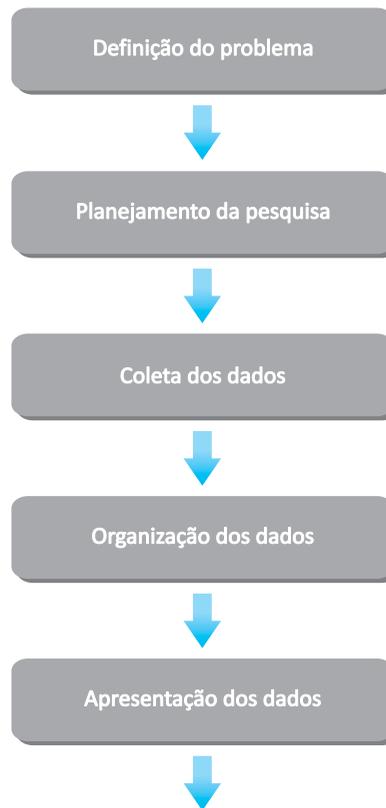


APRESENTAÇÃO DE DADOS

Agora que temos os dados organizados, precisamos apresentá-los e, para tanto, existem duas formas de apresentação que você poderá utilizar, que não se excluem mutuamente, a saber:

- ▶ apresentação por tabelas; e
- ▶ apresentação por gráficos.

Essas formas de expor as informações coletadas permitem sintetizar uma grande quantidade de dados (números), tornando mais fácil a compreensão do atributo em estudo e permitindo uma futura análise.



Na Unidade 2, ampliaremos nossa discussão quanto à forma de apresentação dos dados, ou seja, detalharemos como montar essas tabelas de distribuição de frequências, quais os tipos de gráficos mais adequados para cada situação que você venha a ter.

***Médias** – são os resultados obtidos por meio da soma de todos os valores, divididos pela quantidade de números que você somou. Fonte: Elaborado pelo autor.

ANÁLISE E INTERPRETAÇÃO DE DADOS

Nessa fase, você irá calcular novos números com **médias*** embasadas nos dados coletados. Esses novos números permitem fazer uma descrição do fenômeno em estudo, evidenciando algumas das suas características particulares. Nessa fase, ainda é possível, por vezes, “arriscar” alguma generalização a qual envolverá sempre algum grau de incerteza.

Na fase de análise e de interpretação dos dados, precisamos, ainda, estar muito atentos ao significado das medidas estudadas,

Como quantificar esse grau de incerteza, você irá aprender na Unidade 5.

por exemplo, média e **moda*** e ao porquê de as utilizarmos. Para verificar as relações entre essas medidas, você deve estar de mente aberta. E, para tanto, é necessário que você conheça a estrutura e o cálculo dessas medidas.

***Moda** – valor que mais se repete em um conjunto de observações. Fonte: Elaborado pelo autor.

Imagine que você esteja envolvido em um estabelecimento de conjecturas e na comunicação da informação de uma forma convincente através da elaboração de relatórios, de textos e de artigos que incluam, por exemplo, gráficos e tabelas. As pessoas que se utilizam da estatística como ferramenta devem ser sensibilizadas para a influência que poderá ter o modo de apresentação da informação na comunicação de resultados, a utilização de diferentes gráficos e/ou de diferentes escalas.

Para compreender essa nossa conversa, analise a Figura 1, que apresenta um resumo de todas essas fases:

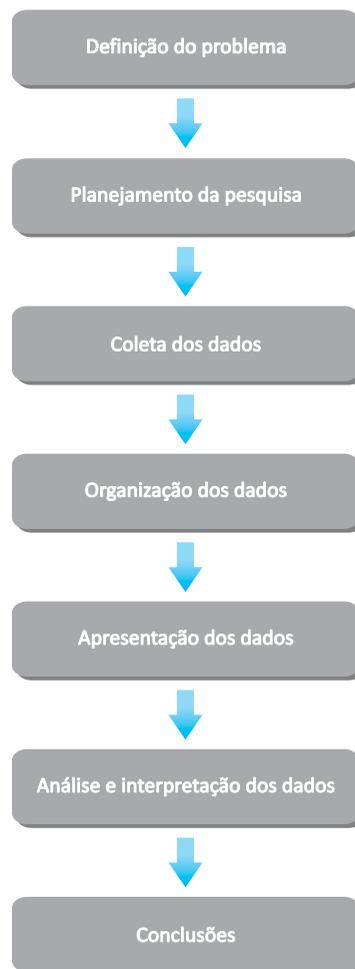


Figura 1: Fases do método estatístico
Fonte: Elaborada pelo autor

Por fim, é importante destacarmos que para a realização dessa fase de análise faz-se necessário que você tenha o domínio da utilização de planilhas tipo Excel e de *softwares* estatísticos.

POPULAÇÃO E AMOSTRAS

Antes, precisamos entender o que é uma população e o que é uma amostra. Se considerarmos somente os habitantes de uma cidade que contribuem com o pagamento do IPTU, podemos considerar essas pessoas como sendo a população, pois apresentam características em comum; sendo, nesse caso, o fato de que elas estão na mesma cidade e contribuem, todas, com o imposto do IPTU.

Suponha, todavia, que você queira trabalhar com apenas uma parte dessa população, a qual apresente as mesmas características da população da qual você terá uma amostra, ou seja, uma porção ou fração da população que preserva todas as características importantes dos elementos que a integram.

Nessa população, geralmente, você poderá medir uma variável, por exemplo, a renda dessas pessoas. Assim, você poderá querer calcular a renda média dessa população de pessoas que contribuem com o IPTU (média populacional (μ) que corresponde, geralmente, a um valor desconhecido chamado de parâmetro). Como você não vai medir toda a população, podemos obter uma amostra que represente essa população e, estudando a amostra, você terá condições de calcular a média amostral (\bar{x}) que corresponde ao estimador, e o resultado obtido (valor numérico) corresponderá à estimativa. Para entender melhor essa relação, observe a Figura 2, a seguir.

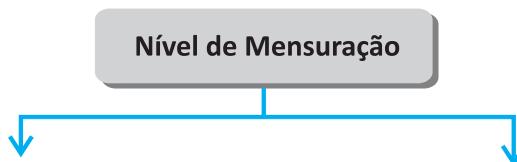


Figura 2: Relações entre estimadores, parâmetros e estimativa
 Fonte: Elaborada pelo autor

Para você entender melhor a figura anterior, verifique que μ (média populacional) e σ (desvio padrão populacional) correspondem aos parâmetros (população), e \bar{x} corresponde ao estimador (amostra) e R\$ 587,00 corresponde à estimativa da renda média populacional (aproximação numérica do valor da população).

Portanto, quando você está estudando uma população inteira (censo) ou realizando uma amostragem, a classificação da variável que está trabalhando será muito importante. Então, vamos ver a classificação das variáveis em relação a sua natureza; as quais podem ser: qualitativas (ordinais ou nominais) e quantitativas (discretas ou contínuas). Essa classificação permitirá, por exemplo, que você defina, posteriormente, o tipo de teste de hipótese a ser utilizado ou o tipo de distribuição de probabilidade que necessitará aplicar para a variável em questão.

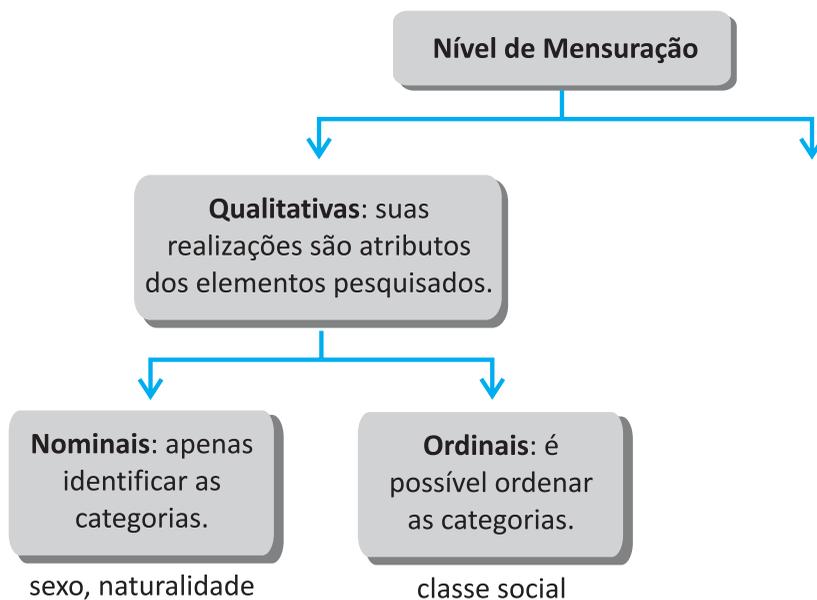
Sendo assim, precisamos entender a classificação das variáveis. Então, mãos à obra! A classificação:



- ▶ **Variável qualitativa:** faz referência a observações relacionadas a atributos que não apresentam estrutura numérica, como cor dos olhos, classe social, estado civil, nome da empresa etc. Essa variável qualitativa pode ser classificada em:
 - ▶ **Nominal:** quando as observações não apresentam nenhuma hierarquia ou ordenamento, como o sexo

dos funcionários de uma prefeitura; estado civil; naturalidade, etc.

- ▶ **Ordinal:** quando as observações apresentam uma hierarquia ou um ordenamento, por exemplo, cargo do funcionário de uma empresa (diretor, gerente, supervisor etc.); posição das empresas em relação a pagamento de impostos em um município (primeira, segunda, terceira etc.).



- ▶ **Variável quantitativa:** essa variável está relacionada às observações que apresentam uma estrutura numérica associada a contagens ou a mensurações, como quantidade de energia elétrica consumida por uma prefeitura em um mês; número de pessoas atendidas por hora em um determinado setor público etc. Essa variável quantitativa pode ser classificada em:
 - ▶ **Discreta:** observações de estrutura numérica estão associadas a valores fixos, ou seja, na maioria dos casos, números inteiros e positivos associados a contagens, como o número de pessoas que pagam seus impostos em dia etc.

- ▶ **Contínua:** são todas as observações que representam valores numéricos que podem assumir qualquer valor dentro de um intervalo, ou seja, os números reais, por exemplo, o tempo que pessoas ficam na fila aguardando para serem atendidas; peso dos funcionários de uma prefeitura etc.

Para melhor visualizar essa classificação das variáveis, observe a Figura 3, apresentada a seguir.

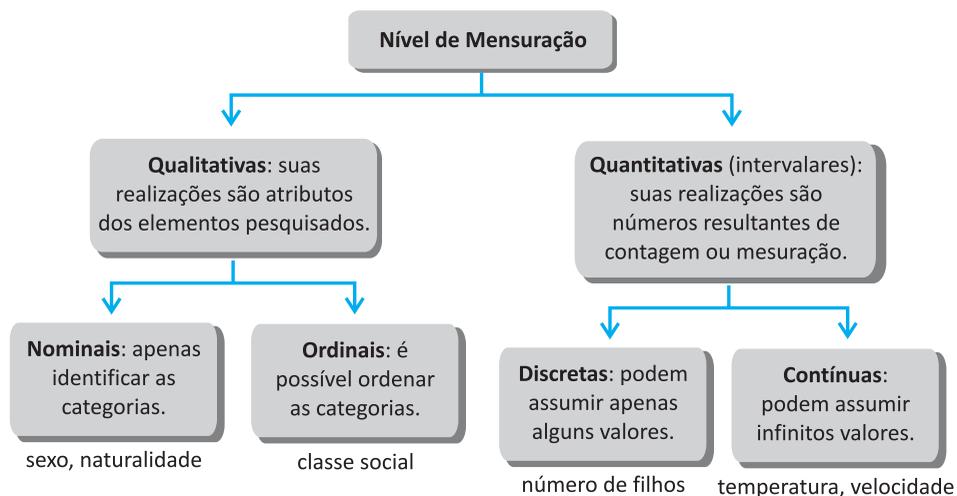


Figura 3: Classificação das variáveis
 Fonte: Elaborada pelo autor

Agora que você já conhece e compreendeu a classificação das variáveis, vamos voltar à relação entre amostragens e populações. A amostragem é o estudo das relações existentes entre a amostra e a população de onde ela – a amostra – foi extraída e a forma como ocorreu essa extração.

As principais vantagens da utilização do estudo por

amostras representativas* em relação ao **censo*** são:

- ▶ Ocorre uma redução no custo, pois sendo os dados obtidos apenas de uma fração da população, as despesas são menores do que as oriundas de um censo. Tratando-se de grandes populações, podemos obter resultados suficientemente precisos, para serem úteis, de amostras que representam apenas uma pequena fração da população.

***Amostras representativas** – são as amostras que mantêm as características da população de onde ela foi retirada. Fonte: Elaborado pelo autor.

***Censo** – avaliação de toda a população. Fonte: Elaborado pelo autor.

- ▶ Na prática ou no dia a dia das organizações, é necessário que os resultados sejam obtidos com a maior rapidez possível. Portanto, com a amostragem, você pode apurar os dados e sintetizá-los mais rapidamente do que em uma contagem completa. Esse é um fator primordial quando se necessita urgentemente das informações. Se o resultado de uma pesquisa for conhecido muito tempo depois, é bem possível que a situação que você pretendia resolver seja, no momento presente, completamente diferente da que existia no momento da coleta dos dados.
- ▶ Outra vantagem corresponde a maior amplitude e flexibilidade. Em certos tipos de investigação, como ocorre em pesquisas de mercado, temos que utilizar pessoal bem treinado e equipamento de alta tecnologia, cuja disponibilidade é limitada para a obtenção de dados. O censo completo torna-se impraticável e resta a escolha de obter as informações por meio de uma amostra. Portanto, com um número reduzido de entrevistadores, por exemplo, o treinamento a ser aplicado com eles é de qualidade muito maior do que em um grupo maior de entrevistadores.
- ▶ A última vantagem a ser citada aqui é a maior exatidão dos resultados. Em virtude de se poder empregar pessoal de melhor qualidade e mais treinado, e por se tornar exequível a supervisão mais cuidadosa do campo de trabalho e do processamento de dados, favorecendo à redução no volume de trabalho, uma amostragem “pode”, na realidade, proporcionar melhores resultados do que o censo.

Dessa forma, podemos dizer que as amostras a serem trabalhadas devem apresentar uma característica importante que corresponde à **representatividade**. Para que as conclusões da teoria de amostragem sejam válidas, as amostras devem ser escolhidas de modo a serem representativas da população.

Antes de darmos continuidade, reflita: como você faria para retirar uma amostra de 300 pessoas que estão em um cadastro de prefeitura que tem 60.000 pessoas? Essa amostra seria representativa da população?

***Plano de amostragem** – plano de como será feita a retirada da amostra da população. Fonte: Elaborado pelo autor.

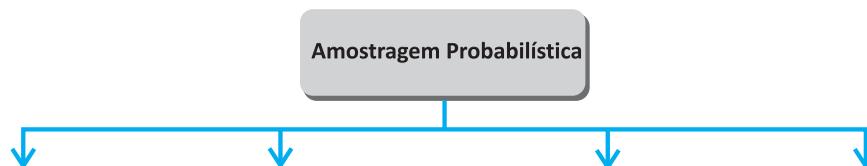
***Unidades amostrais** – correspondem às unidades selecionadas. Fonte: Elaborado pelo autor.

Uma vez que você tenha decidido realizar a pesquisa selecionando uma amostra da população, é preciso elaborar o **plano de amostragem***. O plano de amostragem consiste em definir as **unidades amostrais***, maneira pela qual a amostra será retirada (o tipo de amostragem), e o próprio tamanho da amostra.

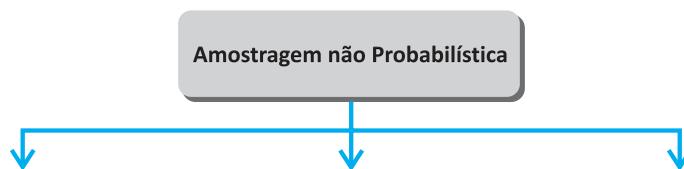
Essas unidades amostrais podem corresponder aos próprios elementos da população, quando há acesso direto a eles, ou qualquer outra unidade que possibilite chegar até eles. Você pode considerar como população os domicílios de uma cidade da qual se deseja avaliar o perfil sócioeconômico. A unidade amostral será cada um dos domicílios, que corresponderá aos elementos da população. Caso a unidade amostral seja definida como os bairros, a unidade amostral não corresponderá aos elementos populacionais.

Temos dois tipos principais de amostragem as probabilísticas e as não probabilísticas. Observe a descrição a seguir:

- ▶ **Amostragem probabilística:** ocorre quando todos os elementos da população tiveram uma probabilidade ou a chance conhecida e diferente de zero de pertencer à amostra. Por exemplo, imagine que temos 50 funcionários de uma prefeitura em uma atividade de treinamento e você deve selecionar 10 funcionários. A realização desse tipo de amostragem somente é possível se a população for finita e totalmente acessível.



- ▶ **Amostragem não probabilística:** é assim denominada sempre que não conhecemos a probabilidade ou a chance de um elemento da população pertencer à amostra. Por exemplo, quando somos obrigados a colher a amostra na parte da população a que temos acesso, ou seja, os elementos da população a que não temos acesso não têm chance de serem sorteados para compor a amostra.



Você pode notar que a utilização de uma amostra probabilística é melhor para garantir a representatividade da amostra, pois o acaso será o único responsável por eventuais discrepâncias entre população e amostra. Essas discrepâncias são levadas em consideração nas inferências estatísticas.

Vamos, então, detalhar os tipos de amostragens probabilísticas.

AMOSTRAGENS PROBABILÍSTICAS

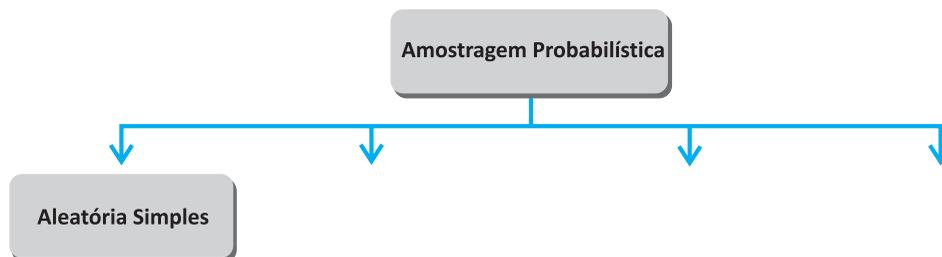
Como já mencionamos, essa amostragem é caracterizada pela chance conhecida de mensurarmos uma amostra. Os principais métodos de amostragem são: aleatória (casual) simples, sistemática, estratificada e conglomerado. Veja a seguir a descrição de cada uma delas.

Amostragem Aleatória (Casual) Simples

Devemos utilizar a Amostragem Aleatória Simples (AAS) somente quando a população for homogênea em relação à variável que se deseja estudar. Geralmente, atribuímos uma numeração a cada indivíduo da população e através de um **sorteio aleatório** os

elementos que irão compor a amostra são selecionados. Todos os elementos da população têm a mesma probabilidade de pertencer à amostra.

Imagine que você queira amostrar um número de pessoas que estão fazendo um determinado concurso com $N = 10.000$ inscritos. Como a população é finita, devemos enumerar cada um dos N candidatos e sortear $n = 1.000$ deles.



Amostragem Sistemática

Em algumas situações como amostrar pessoas que ficam em uma fila, é conveniente retirar da população os elementos que irão compor a amostra de forma cíclica (em períodos), por exemplo, quando os elementos da população se apresentam ordenados. Porém, é de fundamental importância que a variável de interesse não apresente ciclos de variação coincidente com os ciclos de retirada, pois esse fato tornará a amostragem não representativa. Essa técnica de amostragem é o que denominamos de amostragem sistemática.

Para entender melhor, vamos imaginar que você queira retirar uma amostra dos currículos apresentados pelos candidatos em um processo seletivo, e a variável de interesse corresponde à idade deles. Pode ocorrer que pessoas de uma determinada faixa etária deixem para entregar o currículo no último dia. Então, se pegássemos de forma aleatória, poderíamos estar subestimando ou superestimando a idade média. Nessa situação, foram recebidos 500 currículos ordenados por ordem de entrega. Considerando que amostrar 50 currículos é o suficiente para estimar a idade média dos candidatos, utilizamos a técnica de amostragem sistemática, pois pode ocorrer que um grupo de pessoas da mesma faixa etária tenha feito a inscrição em grupo e,

assim, na ordem de inscrição, teremos diversas pessoas com a mesma idade. Devemos considerar então que as idades estejam aleatoriamente distribuídas na população, ou seja, sem qualquer ciclo de repetição.

Para tanto, é necessário, antes, que enumeremos a população de 1 a 500 e calcularemos uma constante (K) que servirá como fator de ciclo para retirada dos currículos amostrados. Assim, podemos dividir os 500 currículos pelo tamanho da amostra (50) que desejamos trabalhar e, então, teremos uma constante igual a 10 e os elementos serão amostrados a cada 10 elementos. Generalizando, teremos que a constante (K) será dada por $K = N/n$, em que N é o tamanho da população e n o tamanho da amostra.

Após a definição do valor de K , fazemos o sorteio de um ponto inicial da amostragem (PIA), ou seja, um dos elementos do primeiro intervalo constituído pelos elementos populacionais numerados de 1 até 10. Na sequência, devemos escolher o próximo que será o elemento de ordem $(i + K)$ e assim por diante, sempre somando K à ordem do elemento anterior até completar a escolha dos n elementos que irão compor a amostra. Um esquema é apresentado na Figura 3 no caso em que $K = 5$.

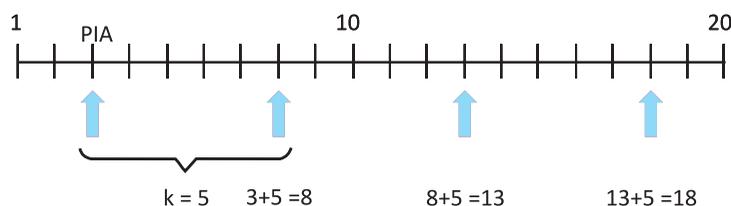


Figura 4: Exemplo de amostra sistemática

Fonte: Elaborada pelo autor

Para fixar os conceitos de amostragem sistemática, vamos fazer, juntos, um esquema de amostragem para saber a opinião dos usuários de um banco, em relação ao tempo de atendimento.

Imagine um Banco X com uma listagem de 33.400 clientes em uma determinada cidade. A pesquisa será feita por telefone, utilizando uma estrutura de call center. Desejando que seja trabalhado com uma amostra de 300 clientes. Como seria organizada a amostragem sistemática?

Antes, você deve dividir o número total de clientes, 33.400, por 300, que é o tamanho da amostra.

$$K = \frac{N}{n} = \frac{33400}{300} = 111,33$$

Como encontramos um valor com casas decimais, então, você irá utilizar um K de aproximadamente 111.

Agora, do primeiro cliente da lista até o de número 111, você irá sortear um número. Vamos considerar que sorteamos o cliente número 10.

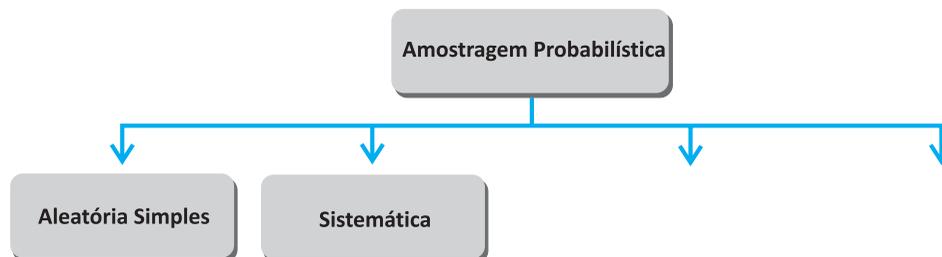
Logo, esse será o primeiro elemento da amostra.

O próximo elemento da amostra será dado pela soma do primeiro sorteado (10º cliente) ao valor de K (111).

Então, o próximo cliente sorteado será o 121º cliente (10 + 111).

Para o sorteio do próximo cliente que irá compor a amostra, teremos o 121º cliente mais o valor de K = 111, ou seja, o 232º cliente.

E, desse modo, você continua até que obtenha todos os elementos da amostra (n = 300 clientes).



Amostragem Estratificada

Quando a variável de interesse apresenta uma heterogeneidade na população e essa heterogeneidade permite a identificação de grupos homogêneos, você pode dividir a população em grupos (estratos) e fazer uma amostragem dentro de cada estrato, garantindo, assim, a representatividade de cada estrato na amostra.

Podemos verificar que pesquisas eleitorais apresentam uma grande heterogeneidade em relação à intenção de votos quando

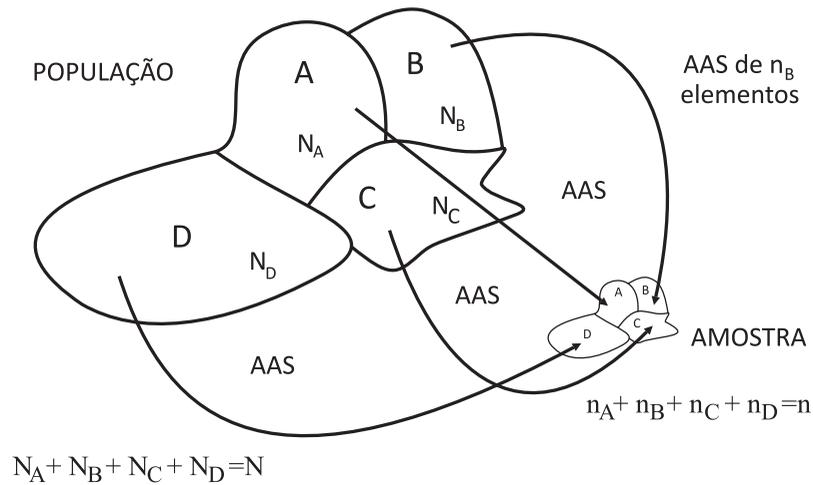
consideramos, por exemplo, a faixa salarial ou o nível de escolaridade. Então, se fizéssemos uma AAS, poderíamos incluir na amostra maior quantidade de elementos de um grupo, embora, proporcionalmente, esse grupo seja pequeno em relação à população. Dessa forma, não teríamos uma amostra representativa da população a ser estudada. Portanto, podemos dividir a população em grupos (estratos) que são homogêneos para a característica que estamos avaliando, ou seja, nesse caso a intenção de votos.

Como estamos dividindo a população em estratos (grupos) que são homogêneos dentro de si, podemos caracterizar a amostragem estratificada. Para efetuarmos a amostragem estratificada de forma proporcional, precisamos, primeiramente, definir a **proporção do estrato em relação à população**.

A proporção do estrato h será igual ao número de elementos presentes nesse estrato (N_h) dividido pelo tamanho da população (N) $\rightarrow (N_h/N)$.

Após você obter essa proporção do estrato em relação à população, deve multiplicar o tamanho total da amostra (n) pela proporção de cada estrato na população (N_h/N).

Dessa maneira, teremos um tamanho de amostra em cada estrato proporcional ao tamanho do estrato em relação à população. A Figura 4 mostra uma população dividida em estratos (grupos) e como é feita a escolha dos elementos de cada estrato (A, B, C, D). Logo, dentro de cada um dos estratos, você pode fazer amostragem usando AAS devido aos estratos serem homogêneos individualmente, considerando a variável de interesse.



Perceba que a quantidade de elementos que iremos sortear dentro de cada estrato é proporcional ao tamanho de cada estrato na população, pois o desenho da amostra é o mesmo da população, porém menor, já que você irá pegar somente uma parte de cada estrato para compor a amostra final.

Para você fixar melhor os conceitos de amostragem estratificada, vamos resolver juntos a seguinte situação: imagine que o governo federal deseja fazer uma pesquisa de satisfação das pessoas em relação a serviços prestados por prefeituras. Estudos anteriores mostram uma relação entre a satisfação das pessoas e o tamanho da cidade. A população a ser considerada diz respeito às cidades de um determinado Estado. Essas cidades foram divididas em três grupos (estratos) levando em conta o tamanho da cidade (pequena, média e de grande porte). Considere que vamos trabalhar com uma amostra de tamanho $n = 200$ cidades e, com as informações a seguir, faça o esquema de uma amostragem estratificada.

ESTRATOS	TAMANHO DO ESTRATO (N° DE CIDADES)
Pequeno porte	$N_1 = 700$
Médio porte	$N_2 = 100$
Grande porte	$N_3 = 27$

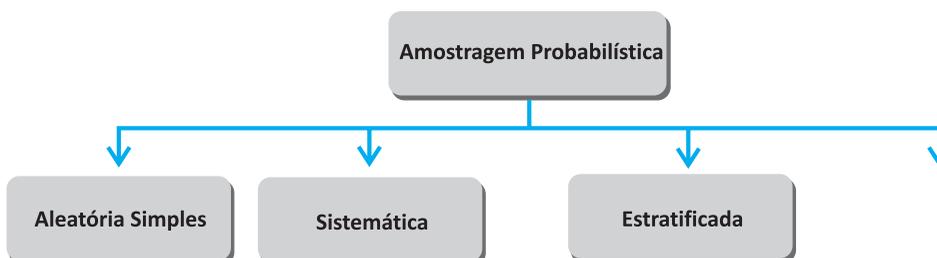
Calcule, antes, a proporção de cada estrato na população, dividindo o tamanho do estrato pelo tamanho da população ($700 + 100 + 27 = 827$).

ESTRATOS	Nº DE CIDADES	PROPORÇÃO
Pequeno porte	$N_1 = 700$	$\frac{N_1}{N} = \frac{700}{827} = 0,8464$
Médio porte	$N_2 = 100$	$\frac{N_2}{N} = \frac{100}{827} = 0,1209$
Grande porte	$N_3 = 27$	$\frac{N_3}{N} = \frac{27}{827} = 0,0326$

A obtenção da quantidade de cidades que será amostrada na população será dada por meio da proporção de cada estrato multiplicado pelo tamanho total da amostra ($n=200$), como é mostrado a seguir:

ESTRATOS	Nº DE CIDADES	PROPORÇÃO	Nº DE CIDADES AMOSTRADA EM CADA ESTRATO
Pequeno porte	$N_1 = 700$	$\frac{N_1}{N} = \frac{700}{827} = 0,8464$	$n_1 = 0,8464 \cdot 200 = 169,28 \cong 170$
Médio porte	$N_2 = 100$	$\frac{N_2}{N} = \frac{100}{827} = 0,1209$	$n_2 = 0,1209 \cdot 200 = 24,18 \cong 24$
Grande porte	$N_3 = 27$	$\frac{N_3}{N} = \frac{27}{827} = 0,0326$	$n_3 = 0,0326 \cdot 200 = 6,52 \cong 6$

Então, na nossa amostra, teremos 170 cidades de porte pequeno, 24 cidades de porte médio e 6 cidades de grande porte.



Amostragem por Conglomerados

Apesar de a amostragem estratificada apresentar resultados satisfatórios, a sua implementação é dificultada pela falta de informações sobre a população para fazer a estratificação.

Para poder contornar esse problema, podemos trabalhar com o esquema de amostragem chamado amostragem por conglomerados.

Os conglomerados são definidos em razão da experiência do gestor ou do pesquisador. Geralmente, podemos definir os conglomerados por fatores geográficos, como bairros e quarteirões. A utilização da amostragem por conglomerados possibilita uma redução significativa do custo no processo de amostragem. Portanto, um conglomerado é um subgrupo da população que, individualmente, reproduz a população. Esse tipo de amostragem é muito útil quando a população é grande, por exemplo, no caso de uma pesquisa em nível nacional.

Você pode estar se perguntando como: realizar uma amostragem por conglomerados?

Apesar de a amostragem por conglomerados, nesse tipo de amostragem, ser utilizada para uma população grande, é simples de calculá-la. Primeiramente, definimos o conglomerado e, assim, dividimos a população nele. Sorteamos os conglomerados por meio de um processo aleatório e avaliamos todos os indivíduos presentes neles, isso é chamado de amostragem por conglomerados em um estágio. Caso façamos um sorteio de elementos dentro de cada conglomerado, teremos uma amostragem por conglomerados em dois estágios. Para entender melhor esse cálculo, observe a Figura 5, que mostra uma amostragem por conglomerados em um único estágio. Cada quadrado corresponde a uma residência. Analise.

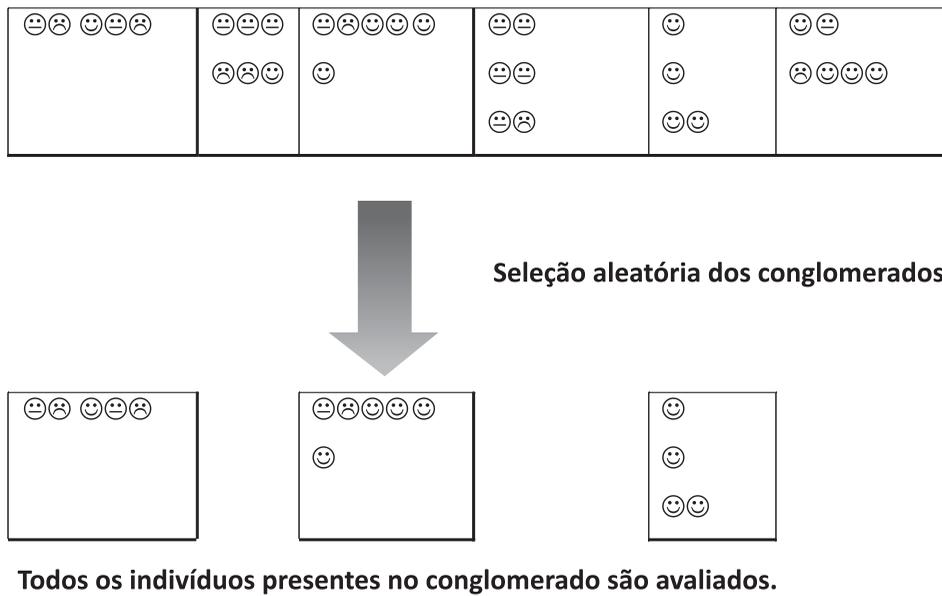
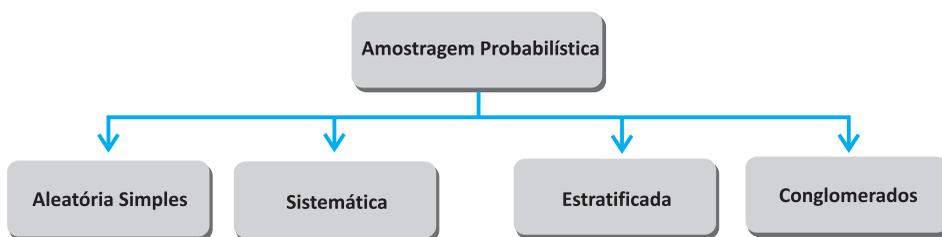


Figura 5: Amostra por conglomerados
 Fonte: Elaborada pelo autor

Um exemplo prático de utilização dessa amostra é a Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE), feita por conglomerados em três estágios.

Para saber mais sobre essa pesquisa acesse www.ibge.com.br.
 Acesso em: 18 nov. 2010.



O cálculo do tamanho amostral será visto em conjunto com a parte de intervalos de confiança na Unidade 5.

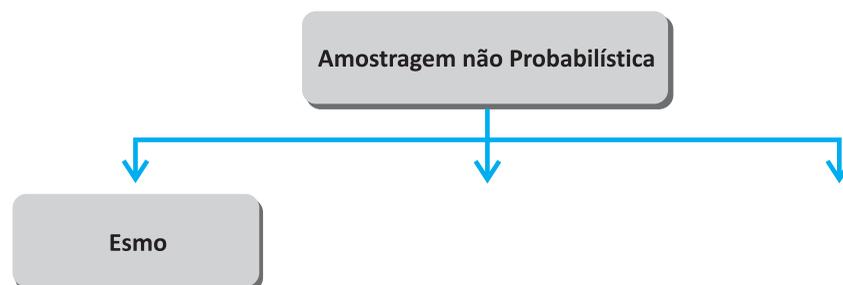
AMOSTRAGEM NÃO PROBABILÍSTICA

Quando trabalhamos com a amostragem não probabilística, não conhecemos *a priori* a probabilidade que um elemento da população tem de pertencer à amostra. Nesse caso, não é possível calcular o erro decorrente da generalização dos resultados das análises estatísticas da amostra para a população de onde a amostra foi retirada. Então, utilizamos geralmente a amostragem não probabilística, por simplicidade ou por impossibilidade de se obter uma amostra probabilística como seria desejável.

Os principais tipos de amostragem não probabilística que temos são: amostragem sem norma, ou a esmo; intencional; e por cotas.

Amostragem a Esmo

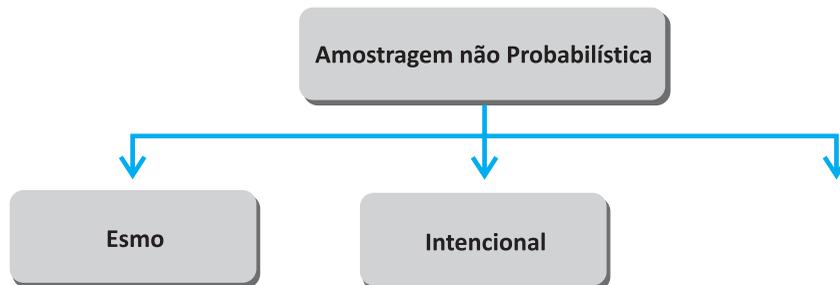
Imagine uma caixa com 1.000 parafusos. A enumeração desses parafusos ficaria muito difícil e a AAS tornar-se-ia inviável. Então, em situações desse tipo, supondo que a população de parafusos seja homogênea, escolhemos a esmo a quantidade relativa ao tamanho da amostra. Quanto mais homogênea for a população, mais podemos supor a equivalência com uma AAS. Dessa forma, os parafusos serão escolhidos para compor a amostra de um determinado tamanho sem nenhuma norma ou a esmo. Daí vem o nome desse tipo de amostragem.



Amostragem Intencional

A amostragem intencional corresponde àquela em que o amostrador deliberadamente escolhe certos elementos para pertencer à amostra por julgar tais elementos bem representativos da população.

Um exemplo desse tipo de amostragem corresponde à situação em que desejamos saber a aceitação de uma nova marca de *whisky* a ser inserida no mercado de uma cidade. Somente entrarão para compor a amostra pessoas que façam uso da bebida e que tenham condições financeiras de comprar essa nova marca (classe social de maior poder aquisitivo).

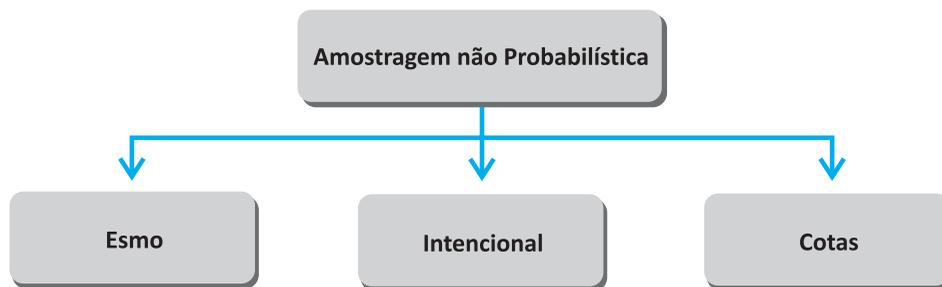


Amostragem por Cotas

Nesse tipo de amostragem, a população é dividida em grupos e, na sequência, é determinada uma cota proporcional ao tamanho de cada grupo. Entretanto, dentro de cada grupo não é feito sorteio, mas sim os elementos são procurados até que a cota de cada grupo seja cumprida.

Encontramos esse tipo de amostra em pesquisas eleitorais quando a **divisão de uma população** ocorre em grupos; considerando, por exemplo, o sexo, o nível de escolaridade, a faixa etária e a renda, que podem servir de base para a definição dos grupos, partindo da suposição de que essas variáveis definem grupos com comportamentos diferenciados no processo eleitoral.

Para termos uma ideia do tamanho desses grupos, podemos recorrer a pesquisas feitas anteriormente pelo IBGE.



Juntando todos os desenhos dos vários tipos de amostragem que fizemos, teremos, então, a Figura 6:

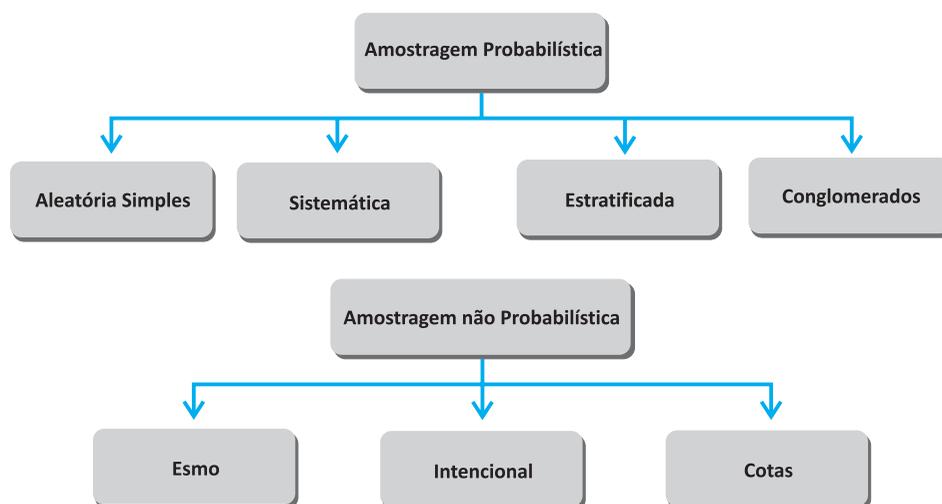


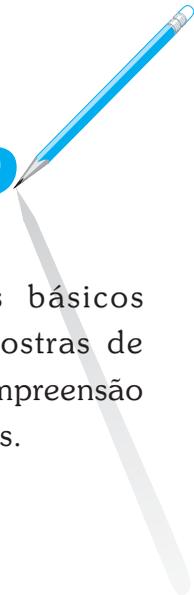
Figura 6: Tipos de amostragem
 Fonte: Elaborada pelo autor

Complementando...

Lembre-se de que a construção do conhecimento é um processo que deve ser cíclico e renovado a cada dia; para tanto, procure descobrir mais acerca desse mundo estatístico seguindo esta orientação:

- 📌 Programa estatístico Bioestat. Disponível em: <<http://www.mamiraua.org.br/download/Default.aspx?dirpath=e:\home\mamiraua\Web\download\BioEstat 5 Portugues&tipo=diretorio>>. Acesso em: 18 nov. 2010. Esse programa permite que você realize os métodos de amostragem, apresentados aqui, computacionalmente.

Resumindo



Nesta Unidade, você conheceu conceitos básicos relacionados à estatística e aprendeu a retirar amostras de populações. Esses conceitos serão importantes para a compreensão de novas informações contidas nas Unidades posteriores.



Atividades de aprendizagem

Depois de ter visto todos os conceitos das fases do método estatístico, a classificação de variáveis e os diferentes planos amostrais, resolva as atividades a seguir. Lembre-se de que as respostas de todas as atividades de aprendizagem estão no final do livro. Em caso de dúvidas, você deve consultar seu tutor.

1. Imagine a situação de um pesquisador que deseje estudar o uso semanal da internet por estudantes de uma escola do Ensino Fundamental. Diferentes perguntas poderiam ser feitas, leia os exemplos e classifique-os em qualitativa nominal ou ordinal e quantitativa discreta ou contínua.
 - a) Você usa internet durante a semana? (sim ou não).
 - b) Qual a intensidade de uso da internet durante a semana? (nenhuma, pequena, média ou grande).
 - c) Quantas vezes você usa a internet durante a semana?
 - d) Por quantas horas você usa a internet durante a semana?
2. Identifique o tipo de amostragem utilizada nas situações a seguir:
 - a) Uma empresa seleciona uma a cada 300 pilhas produzidas em sua linha de produção para a realização de testes de qualidade a fim de conseguir vencer uma licitação pública.

- b) Um pesquisador de empresa aérea seleciona aleatoriamente dez voos para entrevistar todos os passageiros desses voos.
- c) Uma prefeitura testa uma nova estratégia de cobrança selecionando aleatoriamente 250 consumidores com renda inferior a R\$ 300,00 e 250 consumidores com renda de ao menos R\$ 300,00.
- d) Um eleitor indeciso resolve escolher seu candidato da seguinte forma: escreve o nome de cada um deles em cartões separados, mistura-os e extrai um nome, no qual irá votar.
- e) Um pesquisador ficou em um ponto de checagem da polícia (esquina), onde, a cada cinco carros que passavam, era feito um teste de bafômetro para checar a sobriedade do motorista.
- f) Em uma pesquisa com 1.000 pessoas, estas foram selecionadas usando-se um computador para gerar números de telefones para os quais eram, então, discados.
- g) Uma prefeitura, para não perder uma fábrica de montagem de carros, auxiliou em uma pesquisa na qual a empresa dividiu seus carros em cinco categorias: subcompacto, compacto, médio, intermediário e grande; e está entrevistando 200 proprietários de cada categoria para saber da satisfação desses clientes e, assim, ajudar a melhorar as vendas.
- h) Motivada pelo fato de um estudante ter morrido por excesso de bebida, uma universidade fez um estudo sobre o hábito de beber dos estudantes e, para isso, selecionou dez salas de aula e entrevistou os estudantes que lá estavam.