

Livro 1.

A Estatística e o Conhecimento Humano.

1-1. Introdução.

Neste primeiro livro da Coleção “**Aprenda Estatística Através da Pesquisa**”, o prezado leitor será apresentado aos seguintes temas:

- A construção do conhecimento humano: os métodos filosófico e científico;
- A Ciência Estatística. Objeto e objetivos de pesquisa. População e amostra;
- A mensuração da realidade: fenômenos e experimentos aleatórios;
- O espaço amostral de um experimento aleatório;
- Variáveis aleatórias nominais, ordinais, intervalares e proporcionais;
- Variáveis qualitativas discretas: nominais e ordinais;
- Variáveis quantitativas, discretas ou contínuas: intervalares e proporcionais
- Valores amostrais e populacionais de uma variável aleatória;
- Como a estatística descreve a amostra através das “estatísticas descritivas”
- Como a estatística induz conclusões populacionais a partir de valores amostrais;
- A probabilidade como uma previsão do futuro;
- Métodos de amostragem e representatividade da amostra perante a população;
- O método estatístico na construção do conhecimento humano;
- A construção de um quadro conceptual em uma pesquisa estatística.

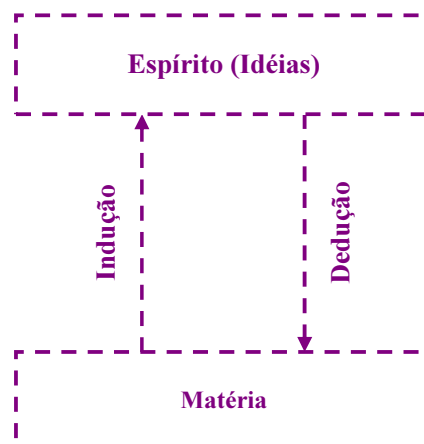
1-2. A Construção do Conhecimento Humano.

Da mesma forma que a **matéria** se transforma em **energia**, transforma-se também em **idéias**. A construção do **conhecimento humano** é o processo no qual a realidade material se transforma em realidade conceitual ou espiritual, é o processo de transmutação do mundo exterior no qual vivemos em idéias no interior do cérebro humano.

A **filosofia** é a ciência que estuda as leis gerais que regem essa transformação, ou seja, estuda como se constróem as representações da realidade objetiva no pensamento humano, representações essas que, por sua vez, passam a fazer parte também da realidade exterior na qual estamos todos imersos corporal e espiritualmente.

O método filosófico é composto pela **indução**, que generaliza a observação de partes de um objeto de estudo para reconstruir o todo desse objeto no plano das idéias, e pela **dedução**, que particulariza a idéia do todo para prognosticar futuros comportamentos materiais. Veja-se o esquema ao lado.

Relacionando matéria e idéia, o método filosófico de indução e dedução é **materialista** quando considera ser a matéria a base fundamental de onde se constróem as idéias, ou é **idealista**, quando supõe



ser a idéia o fundamento que precede e determina o mundo material.

O método filosófico é **dialético**, quando supõe a existência do movimento como condição intrínseca da matéria, o que se reproduz no movimento das idéias, ou é **metafísico**, quando considera que as idéias são estáticas, não possuem movimento. O método filosófico é **indeterminista** quando supõe que o futuro, embora seja uma consequência do presente, não pode ser completamente previsto ou **determinista**, se supõe que todas as relações do universo são causais, sendo o futuro perfeitamente predeterminado.

A **metafísica idealista** supõe a existência de uma idéia primeira e imutável, predecessora do mundo material, o qual, em consequência, deve ser também considerado imutável e determinado. Essa corrente filosófica se identifica com a religião, no sentido amplo da palavra, onde Deus (idéia primeira) criou todo o Universo (mundo material) e detém o conhecimento e o poder sobre todo o futuro, que é de antemão por ele determinado.

O **materialismo dialético** explica que a idéia é uma consequência do mundo material, um produto dos neurônios humanos, uma consequência do esforço que o animal “homo sapiens” realiza para viver e se reproduzir. Mas, sendo consequência da matéria, as idéias possuem movimento, se transformam, geram novas idéias e, justamente por isso, influem decisivamente nas transformações do mundo material. De fato, cada uma das intervenções humanas sobre o Planeta Terra é precedida pela idéia de tal intervenção, dando origem a um infundável ciclo matéria-idéia-matéria. Dado o próprio caráter da matéria em movimento, o futuro pode ser determinado apenas até certo limite, a partir de onde nenhuma previsão totalmente determinista é possível.

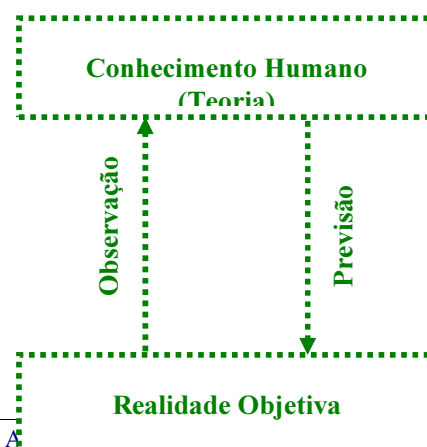
1-3. O Método Científico.

Se estuda as leis gerais da transformação da matéria em espírito, a filosofia não realiza ela própria a criação do conhecimento humano, a qual fica a cargo das diversas ciências. Cada ciência, em seu campo particular de estudo, realiza seu próprio ciclo de indução e dedução. A realidade objetiva (matéria) é transformada em teoria (idéia) através da observação. No caso da observação ocorrer em condições tais que a realidade é controlada pelo observador, dá-se-lhe o nome de experimentação.

A teoria, uma vez formulada, é utilizada para prever o comportamento futuro da realidade, possibilitando, dessa forma, o projeto de novas experimentações, as quais visam encontrar exceções ao comportamento teórico capazes de produzir teorias mais gerais.

A ciência é materialista pois parte da observação do mundo material e é dialética, pois leva em conta o movimento da matéria e das idéias. Considera que o comportamento da realidade objetiva pode ser determinado ou indeterminado, a depender dos limites da observação do fenômeno em estudo.

O método científico não se propõe a encontrar um final, a afirmar um dia que todo o conhecimento do Universo será alcançado. Mas possibilita



declarar que o Universo é **cognoscível** ou passível de vir a ser conhecido cada vez mais e sem limites através de sucessivas transformações matéria-idéia-matéria.

1-4. Objeto e Objetivos de Pesquisa.

Pesquisar é transformar realidade em idéia. O primeiro passo dessa transformação é delimitar os limites da realidade e definir o **objeto** que se quer estudar. O objeto pode ser físico (por exemplo, a tensão de ruptura de um certo tipo de aço), natural (o cultivo de um determinado vegetal de interesse), social (o processo de decisão do eleitorado entre os diversos candidatos à presidência) ou psicológico (a inteligência de uma pessoa). Assim, **o objeto de pesquisa é aquilo que se quer estudar, a parcela do universo físico, natural, social ou psicológico se quer transformar em idéias.**

Os **objetivos da pesquisa são aquilo que se quer saber acerca do objeto pesquisado.** Se o objeto de estudo for um material de construção, um objetivo seria determinar sua tensão de ruptura. Se for o cultivo de certo vegetal, um objetivo seria verificar a quantidade de adubo ótima em dado método de cultivo. Se for a inteligência humana, um objetivo poderia ser a determinação do QI médio de um grupo de estudantes. Se o objeto for o processo de decisão do eleitorado, um objetivo seria determinar qual candidato possui a preferência dos eleitores.

O objetivo da pesquisa pode ser confirmar ou negar uma ou mais **hipóteses de pesquisa.**

Uma hipótese poderia afirmar que dois tipos de metal diferentes possuem a mesma tensão de ruptura. Outra poderia estabelecer que o método de cultivo A é mais produtivo que o método B. Outra ainda poderia considerar que adultos que passaram fome na primeira infância possuem QI mais baixo que adultos bem alimentados nessa fase da vida. Finalmente, uma hipótese poderia afirmar que uma pesquisa eleitoral pode ser utilizada para induzir a opinião dos cidadãos.

Embora a Estatística seja uma ciência com um método próprio, ela atua na produção do conhecimento através de objetos das demais ciências. Diferentes objetos de estudo exigem a aplicação da Estatística por diferentes cientistas. Para se determinar a tensão de ruptura de um metal quem estará utilizando a Estatística será um engenheiro de materiais. Para se verificar qual será a quantidade ótima de adubo, será o agrônomo o usuário. Na estimativa da intenção de voto, será utilizada pelo cientista social. A determinação do QI, exige do psicólogo o conhecimento estatístico.

A Estatística sozinha jamais chegaria a uma conclusão acerca desses objetos de estudo, mas, em contraposição, apresenta um método unificado capaz de se desdobrar e servir de ferramenta para o estudo de todos eles. Tal relação entre ciências se traduz em relações entre seres humanos: como os profissionais de cada uma dessas áreas via de regra não dominam a Estatística e o estatístico não pode se aprofundar em todas elas, a realização de pesquisas exige o **diálogo** entre os diversos pesquisadores e a formação de equipes interdisciplinares em que um profissional se aprofunda no campo do conhecimento do outro, colaborando ambos para o aprofundamento da investigação.

1-5. O que é População?

Digamos que o professor de Estatística tenha como **objeto** a opinião dos estudantes de sua turma de ciências humanas acerca dessa disciplina. Um **objetivo** seria saber em que nível essa população possui bloqueios com matemática. Mas o professor poderia estar interessado em todos os estudantes de ciências humanas aptos a cursar Estatística. Nesse caso, a população seria mais ampla. Ou o poderia estar interessado em conhecer a opinião de todos os alunos da UFBA acerca da Estatística. E se fosse o MEC o interessado, a população seria composta de todos os estudantes universitários brasileiros.

Uma população, em uma pesquisa, é um conjunto que engloba todos os elementos detentores de uma ou mais características em comum que os relacionam ao objeto de estudo. Sua amplitude depende do objeto e dos objetivos pesquisados. Depois de se estabelecer o objeto e os objetivos de estudo, o passo seguinte em uma pesquisa é a delimitação da população dentro da qual se deseja responder tais objetivos com vistas ao estudo de tal objeto. **Delimitar a população é estabelecer o campo de validade do conhecimento que se quer obter.**

Delimitar a população não é tão fácil como pode parecer. A população de certa turma de Estatística é perfeitamente delimitada pela listagem expedida pela Secretaria Geral de Cursos. Se a população forem os estudantes **aptos** a cursarem Estatística, no entanto, em todo final de semestre haverá uma renovação parcial. Somente se delimitássemos um período de tempo seria possível obter a lista de estudantes, mas poderia não ser interessante do ponto de vista do objeto de estudo a delimitação de um período restrito de tempo. Um outro exemplo: quem são precisamente os moradores de um bairro popular? A própria região física de um bairro é muito difícil de se demarcar, pois suas fronteiras estão em permanente movimento e além disso, as pessoas estão sempre se mudando, umas nascem, outras morrem....

Podemos, em vista do exposto, classificar as populações em **finitas**, ou **fechadas**, quando é possível obter uma lista exata dos membros, ou **abertas**, **infinitas**, quando existem processos de renovação que impem o estabelecimento de limites exatos.

1-6. O que é Amostra?

Analisemos a hipótese de que o MEC queira estudar a opinião acerca da Estatística (o objeto da pesquisa) dos estudantes universitários brasileiros (a população). Como faria? Iria pesquisar todos os estudantes do Brasil? Não seria mais econômico tomar um subconjunto de alunos e pesquisá-los, confiando que as conclusões obtidas a partir desse subconjunto seriam válidas para todos os demais estudantes? Claro que a amostra teria que ser suficientemente grande e subdividida através das regiões e universidades para representar toda a população de forma confiável. Mas, de qualquer maneira, isso representaria uma economia considerável em termos de tempo e custos de pesquisa.

O mesmo iria ocorrer se a população pesquisada fossem todos os estudantes da UFBA ou os estudantes de ciências humanas aptos a cursarem Estatística. E mesmo no caso da população ser uma determinada turma de Estatística, talvez nem todos estivessem presentes quando o professor aplicasse um questionário visando coletar a opinião dos estudantes. A amostra, nesse último caso, seriam os alunos e alunas que respondessem ao questionário aplicado.

Amostra, do ponto de vista estatístico, é **qualquer subconjunto da população** do qual se extraem dados numéricos com a finalidade de se obter uma representação dessa população que permita o estudo de certo objeto com a qual ela se relaciona.

1-7. O Problema da Mensuração da Realidade Objetiva.

Chegamos ao começo do estudo da estatística propriamente dita com uma pergunta fundamental. Como os dados numéricos são extraídos da realidade material? Como transformar a realidade material em números? A procura da resposta nos leva ao estudo da **mensuração** da realidade objetiva.

Mensurar a realidade objetiva significa **classificar** seus fenômenos entre **qualidades** ou categorias distintas, bem como **contar** ou **medir** a **quantidade** existente de uma determinada qualidade ou propriedade.

Já vimos que o caráter da matéria é simultaneamente determinado e indeterminado. Todos os fenômenos da natureza têm uma componente **determinística** e outra indeterminística, ou **aleatória**. A importância de cada uma varia de caso para caso, mas as duas estão sempre presentes. Quando um estudo ultrapassa os limites do determinismo, dizemos que estamos diante de um **fenômeno aleatório**.

Toda vez que classificamos ou medimos um fenômeno aleatório realizamos um **experimento aleatório**. Em oposição aos experimentos ditos determinísticos, que realizados em igualdade de condições levam sempre a resultados iguais, um experimento aleatório, repetido sob condições as mais idênticas possíveis, leva em geral a resultados diferentes.

Experimentos totalmente deterministas são, na verdade, idealizações do cérebro humano, utilizadas para se obter uma primeira aproximação do conhecimento sobre a natureza. Como a matéria é simultaneamente determinada e indeterminada, os experimentos ditos deterministas somente o são dentro de certos limites de observação. Toda vez que ultrapassarmos esses limites, estaremos diante de um experimento aleatório.

Um corpo que cai segue as leis da mecânica, mas tais leis só determinariam completamente sua trajetória se tal corpo pudesse cair no vácuo perfeito. Na vida real, um corpo que cai estará submetido à resistência do ar, a qual é função da condição meteorológica no momento da queda. Junto com a lei da gravidade, haverá também uma ação **aleatória** do ar sobre o corpo e, em certas situações, como no salto de um pára-quedista, sua trajetória não poderá ser determinada exatamente.

Medir o peso de um saco de **1kg** de açúcar em uma balança de supermercado é um experimento que o senso comum julga ser apenas determinístico, mas que possui uma componente aleatória. O prezado leitor que verifique com os próprios olhos!

Determinar a exata posição de uma estrela no céu é um experimento aparentemente determinístico que, ao ser analisado pela teoria dos erros de observação, revela uma componente aleatória. De fato, foi estudando tais erros de observação que Gauss desenvolveu a chamada “Curva de Gauss”, ou a equação do modelo normal de probabilidades, um dos mais utilizados pela Estatística.

A Estatística se ocupa de examinar os fenômenos da natureza dentro dos limites de seu indeterminismo, ou aleatoriedade, ou seja, o trabalho estatístico se inicia no ato de **mensurar** (classificar ou medir) fenômenos onde o **caráter aleatório** é predominante.

1-8. Fenômenos e Experimentos Aleatórios

A seguir, discutiremos alguns exemplos de tais fenômenos e experimentos aleatórios e veremos como seus resultados podem ser mensurados.

O **sexo** é um fenômeno aleatório. Nascer homem ou mulher só depende da tranza entre os espermatozoides e o óvulo, cujo resultado a ninguém é dado saber por antecipação. Verificar o sexo de uma pessoa previamente escolhida por sorteio é um experimento aleatório. Anteriormente ao sorteio, todas as pessoas são iguais, independente do sexo de cada um, visto ser impossível saber de que sexo será o próximo a ser sorteado. Após ter sido escolhida, cada pessoa revelará seu sexo, o qual não necessariamente será igual ao da pessoa anteriormente escolhida.

A exata **instrução formal** de cada pessoa é um fenômeno aleatório. Quem tem a sorte de não passar fome na primeira infância possui inteligência para estudar, e a vida concreta condiciona o restante. Digamos que o professor queira estudar a população de estudantes de sua turma, no tocante ao tipo de ensino secundário cursado. É um experimento aleatório o ato de solicitar em um questionário que um estudante responda a seguinte pergunta:

Você fez o 2º Grau em que tipo de escola?

- ☐ *Só em Pública Estadual*
- ☐ *Maior Parte na Pública Estadual*
- ☐ *Só em Pública CEFET*
- ☐ *Maior Parte na Pública CEFET*
- ☐ *Só na Particular*
- ☐ *Maior Parte na Particular*

Antes de receber a lista de nomes de sua disciplina, o professor nada conhece acerca do histórico de instrução formal dos estudantes que serão entrevistados. Cada um possui seu- histórico, que se revela depois da aplicação do questionário. A pergunta é sempre a mesma e os estudantes tem todos em comum o fato de estarem matriculados na disciplina. As respostas, no entanto, são em geral diferentes de um estudante para outro, não podendo serem previstas com antecipação nem existindo qualquer determinação da resposta de um estudante em função da resposta de outro.

Sendo a história de instrução formal de uma pessoa um fenômeno aleatório, é também um experimento aleatório solicitar a um estudante que responda a uma das questões definidoras do **nível sócio-econômico** do entrevistado segundo o “Critério Brasil” adotado pelo IBOPE. Por exemplo, seja a questão seguinte:

Marque o grau de instrução do Chefe de Família (sic) ⁽¹⁾ em sua casa:

¹ O reacionário conceito de “Chefe de Família”, abolido do Código Civil Brasileiro, continua presente na definição do índice sócio econômico do instituto de opinião vinculado à Rede Globo....

- ☐ Analfabeto/Primário incompleto
- ☐ Primário completo/Ginásial incompleto
- ☐ Ginásial completo/Colegial incompleto
- ☐ Colegial completo/Superior incompleto
- ☐ Superior completo

A **opinião** de uma pessoa é um fenômeno aleatório. Mesmo sendo condicionada pela mídia, cada pessoa tem sua experiência concreta e certa massa cinzenta. Pensa, uma vez que existe enquanto ser humano. Em consequência, a obtenção de **respostas** a cada uma das perguntas de um **questionário fechado** é um experimento aleatório. As perguntas são sempre as mesmas e todos os entrevistadores são treinados para procederem de forma padronizada perante os entrevistados. Estes, no entanto, respondem cada qual expressando sua opinião, gerando resultados em geral diferentes.

Um questionário fechado que peça aos entrevistados responderem “**sim**” ou “**não**”, caso considerem que determinado político possua ou não possua cada uma dentre 12 qualidades de uma lista é um experimento aleatório:

Marque abaixo **sim** ou **não**, conforme você considere que FHC possua as seguintes qualidades:

- | | | |
|---|------------------------------|------------------------------|
| (1) É honesto | ☐ Sim | ☐ Não |
| (2) É sincero | ☐ Sim | ☐ Não |
| (3) É trabalhador..... | ☐ Sim | ☐ Não |
| (4) É responsável | ☐ Sim | ☐ Não |
| (5) É fiel aos compromissos que assume..... | ☐ Sim | ☐ Não |
| (6) É patriota | ☐ Sim | ☐ Não |
| (7) É inteligente | ☐ Sim | ☐ Não |
| (8) É culto | ☐ Sim | ☐ Não |
| (9) É preparado..... | ☐ Sim | ☐ Não |
| (10) É corajoso..... | ☐ Sim | ☐ Não |
| (11) Possui visão do futuro..... | ☐ Sim | ☐ Não |
| (12) É comprometido com o povo..... | ☐ Sim | ☐ Não |
| Total de qualidades marcadas como “Sim” = | | |
| Total de qualidades marcadas como “Não” = | | |

A **transmissão do conhecimento humano** é um fenômeno aleatório. Submeter estudantes a uma prova é um experimento aleatório. As questões são as mesmas, o tempo de resposta é igual e a todos foram oferecidas as mesmas aulas, mas as notas são em geral diferentes. Uma prova pode ser oral, escrita ou prática. Uma prova prática é também um experimento aleatório.

Suponha que o professor queira pesquisar o **conhecimento de informática** dos estudantes e aplique uma prova prática, onde cada estudante é solicitado a executar em um computador **20** tarefas. Acompanhando o teste, um monitor anota o acerto ou o erro em cada questão. Embora o computador seja o mesmo, as questões iguais e o monitor treinado para agir de forma padronizada, o número de acertos será em geral diferente de um estudante para outro, não podendo ser previsto com antecipação.

A **temperatura corporal** de uma pessoa, sã ou em estado de doença, é um fenômeno aleatório. Medir com um termômetro de décimo de grau a temperatura corporal de um paciente num hospital é um experimento aleatório. O termômetro é o mesmo e o auxiliar de enfermagem executa um procedimento padronizado, mas os resultados das medidas são em geral diferentes. Mesmo as leituras observadas em um mesmo paciente sob tra-

tamento possuem um comportamento aleatório, quando lidas com um instrumento suficientemente preciso.

A **movimentação das pessoas em uma cidade** é um fenômeno aleatório. Entrar em um ônibus e contar o **número de passageiros** é um experimento aleatório. Os ônibus são iguais no sentido de todos serem veículos que transportam pessoas coletivamente. O faixa de horário é a mesma, por exemplo, no horário de pico ao final da tarde ou ao se iniciar a manhã. As pessoas se igualam no sentido de todas serem usuárias dos sistema de transporte coletivo. No entanto, o número de passageiros contados em cada ônibus resultará em geral diferente, não podendo ser previsto por antecipação quantos passageiros estará levando o próximo ônibus a virar a esquina e encostar no ponto.

A **altura** de uma pessoa é um fenômeno aleatório, dependendo de fatores como hereditariedade, tipo de alimentação, hábitos esportivos e, em consequência, nível sócio-econômico. Medir com uma régua centimétrica a altura de uma estudante escolhida para fazer parte de uma amostra é um experimento aleatório. Apesar da utilização da mesma régua e técnica de medida o resultado será diferente para cada estudante. Como é óbvio, poderá haver resultados iguais, dado que a escala de medida da régua centimétrica possui um número limitado de graduações. Mas os valores da altura de uma estudante não são determinados, não são previsíveis por antecipação, irão aparecer em uma pesquisa segundo as regras do incerto que governaram a escolha dos elementos da amostra.

1-9. O Espaço Amostral de um Experimento Aleatório.

Como classificar ou mensurar um fenômeno aleatório é sempre um experimento aleatório e os resultados de tais experimentos não podem ser previstos com antecipação, definimos para cada experimento aleatório um conjunto que contenha **todos os possíveis resultados desse experimento**, dando a ele o nome de **espaço amostral**. O espaço amostral contém todos os resultados de um experimento, mas não precisa se restringir a apenas tais resultados: podem “sobrar” resultados que nunca irão ocorrer. O que não pode é “faltar” resultados possíveis.

O **espaço amostral** não deve ser confundido com a **amostra**. É o espaço utilizado para **medir** ou para **classificar** os elementos da amostra, mas não é a amostra ela mesma. A diferença entre espaço amostral e amostra fica clara quando se nota que **todos os resultados possíveis do experimento aleatório devem pertencer ao espaço amostral**, e não apenas os que efetivamente ocorreram durante as repetições desse experimento realizadas com vistas a se coletar os dados da amostra.

Cada elemento da amostra gera uma repetição do experimento aleatório, mas a amostra não necessariamente exaure todos os casos possíveis do experimento. Se o espaço amostral tiver muitos casos possíveis, inúmeros pontos amostrais poderão não ocorrer durante as repetições realizadas para a coleta de dados da amostra. Nesse caso, sobrarão pontos amostrais que não ocorreram na amostragem, mas, pela própria definição de espaço amostral, não poderão faltar pontos amostrais adequados a descrever cada resultado obtido na amostra.

É possível elaborar mais do que um conjunto que contenha todos os possíveis resultados de um experimento aleatório: um experimento aleatório pode possuir diferentes espaços amostrais, de acordo com a conveniência da análise que se irá realizar.

Veremos a seguir alguns espaços amostrais que podem ser associados aos diversos experimentos aleatórios que estudamos até o momento. Um espaço amostral definido pelo experimento de classificar, em função do sexo, uma pessoa sorteada é:

$$\Omega_1 = \text{Sexos} = \{\text{O Sexo da Pessoa é Feminino; O Sexo da Pessoa é Masculino}\}$$

Um espaço amostral para o experimento de solicitar a um estudante que responda o tipo de ensino secundário por ele cursado pode ser o seguinte:

$$\Omega_2 = \left\{ \begin{array}{l} \text{Fez o 2º Grau só em Escola Pública Estadual} \\ \text{Fez o 2º Grau só em Escola Pública CEFET} \\ \text{Fez a maior parte do 2º Grau em Escola Pública Estadual} \\ \text{Fez a maior parte do 2º Grau em Escola Pública CEFET} \\ \text{Fez a maior parte do 2º Grau em Escola Particular} \\ \text{Fez o 2º Grau só em Escola Particular} \end{array} \right\}$$

Um espaço amostral para o experimento de solicitar o grau de instrução do “Chefe de Família” poderia ser:

$$\Omega_3 = \left\{ \begin{array}{l} \text{“O chefe de família é Analfabeto ou cursou incompletamente o Primário”} \\ \text{“O chefe de família completou o curso o Primário ou cursou incompletamente o Ginásial”} \\ \text{“O chefe de família completou o curso Ginásial ou cursou incompletamente o Colegial”} \\ \text{“O chefe de família completou o curso Colegial ou cursou incompletamente o Superior”} \\ \text{“O chefe de família completou o curso Superior”} \end{array} \right\}$$

Um espaço amostral para o experimento definido pela questão fechada que pede aos entrevistados responderem “sim” ou “não”, caso considerem que determinado político possua ou não possua cada uma dentre 12 qualidades de uma lista poderia ser:

$$\Omega_4 = \left\{ \begin{array}{l} \text{“O entrevistado respondeu } S_1 S_2 S_3 S_4 S_5 S_6 S_7 S_8 S_9 S_{10} S_{11} S_{12} \text{”} \\ \text{“O entrevistado respondeu } N_1 S_2 S_3 S_4 S_5 S_6 S_7 S_8 S_9 S_{10} S_{11} S_{12} \text{”} \\ \dots \dots \dots \dots \dots \\ \text{“O entrevistado respondeu } N_1 N_2 N_3 N_4 N_5 N_6 N_7 N_8 N_9 N_{10} N_{11} N_{12} \text{”} \end{array} \right\}$$

Um espaço amostral para um teste prático com 20 questões de informática, onde cada questão é considerada certa ou errada, poderia ser o seguinte conjunto:

$$\Omega_5 = \left\{ \begin{array}{l} \text{CCCCCCCCCCCCCCCCCCCC} \\ \text{ECCCCCCCCCCCCCCCCCCC} \\ \text{CECCCCCCCCCCCCCCCC} \\ \dots \\ \text{CCCCCCCCCCCCCCCCCCCC} \\ \text{EECCCCCCCCCCCCCCCC} \\ \text{ECECCCCCCCCCCCCCCCC} \\ \text{ECCECCCCCCCCCCCCCCCC} \\ \dots \\ \text{EEEEEEEEEEEEEEEEEEEE} \end{array} \right\}$$

Um espaço amostral definido pelo experimento de se ler em um termômetro digital de décimo de grau a temperatura de uma pessoa pode ser:

$$\Omega_{6A} = \left\{ \begin{array}{c} \dots \dots \dots \\ \text{"A temperatura medida foi 30 graus e 0 décimos"} \\ \text{"A temperatura medida foi 30 graus e 1 décimos"} \\ \text{"A temperatura medida foi 30 graus e 2 décimos"} \\ \dots \dots \dots \\ \text{"A temperatura medida foi 44 graus e 9 décimos"} \\ \dots \dots \dots \end{array} \right\}$$

Um espaço amostral definido pelo experimento de se contar o número de passageiros em um ônibus em determinado horário pode ser:

$$\Omega_{7A} = \left\{ \begin{array}{c} \text{"Havia 0 passageiros no ônibus"} \\ \text{"Havia 1 passageiro no ônibus"} \\ \text{"Havia 2 passageiros no ônibus"} \\ \dots \dots \dots \\ \text{"Havia 107 passageiros no ônibus"} \\ \text{"Havia 108 passageiros no ônibus"} \\ \dots \dots \dots \end{array} \right\}$$

Um espaço amostral definido pelo experimento de se medir com régua centimétrica a altura de uma estudante poderia ser:

$$\Omega_{8A} = \left\{ \begin{array}{c} \dots \dots \dots \\ \text{"A altura medida foi de 1 metro e 39 centímetros"} \\ \text{"A altura medida foi de 1 metro e 40 centímetros"} \\ \text{"A altura medida foi de 1 metro e 41 centímetros"} \\ \dots \dots \dots \\ \text{"A altura medida foi de 2 metros e 0 centímetros"} \\ \text{"A altura medida foi de 2 metros e 1 centímetros"} \\ \dots \dots \dots \end{array} \right\}$$

1-10. Variáveis Aleatórias.

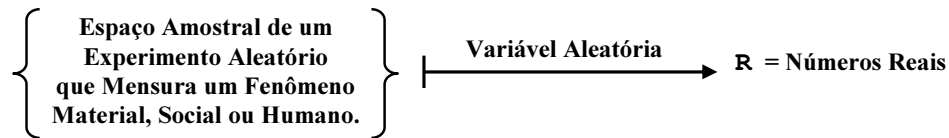
Procurando mensurar os fenômenos aleatórios, estudamos o que são experimentos aleatórios e definimos o que vem a ser espaço amostral. Estudaremos agora a função matemática que finaliza o processo de transformação da realidade objetiva em números, a qual é chamada de **variável aleatória**. Tal conceito é a peça-chave de um estudo estatístico, visto que as variáveis aleatórias são utilizadas para representar as características que se quer estudar em determinada população.

Definimos uma variável aleatória como uma **função** cujo **domínio** é um **espaço amostral** e cuja **imagem** está contida no conjunto dos **números reais**. Em linguagem matemática:

$$X: \Omega \rightarrow \text{Imagem}(X) \subset \mathbb{R}$$

O domínio de uma função é a região onde ela atua, ou de onde ela parte, e a imagem é o conjunto que a função univocamente associa a cada elemento do domínio. Tendo por domínio um espaço amostral, delineado em função de um experimento aleatório de

mensurar (medir ou classificar) a realidade objetiva, e por imagem os números reais, uma variável aleatória estabelece uma relação funcional entre esses dois conjuntos e expressa numericamente eventos da realidade física, natural, social ou humana:



É sempre possível atribuir um **modelo de distribuição populacional de probabilidades** a cada variável aleatória. A importância disso está em a permitir o cálculo de probabilidades de eventos populacionais nos quais estamos interessados. Tais modelos nada mais são que variáveis aleatórias definidas em função de experimentos aleatórios idealizados de forma tal a permitirem o cálculo teórico de sua distribuição de probabilidades. As variáveis **Binomial** e de **Poisson**, ambas discretas, **Normal**, **Student**, **Qui-Quadrado**, **F** ou de **Fisher**, contínuas, são alguns modelos bastante utilizados

Veremos a seguir as variáveis aleatórias que podem ser associadas aos espaços amostrais dos experimentos aleatórios que estivemos debatendo.

O experimento aleatório de se perguntar em um questionário o sexo de cada estudante escolhido para compor determinada amostra dá origem à variável aleatória:

$$X_1 = \text{Sexo de um Estudante.}$$

que associa a cada ponto do espaço amostral um elemento do conjunto:

$$\text{Imagem}(X) = \{1, 0\}$$

de acordo com a seguinte regra:

Sexo do Estudante	Valor de X_1
Feminino	1
Masculino	0

O caráter aleatório dessa variável é dado pelo processo de escolha da amostra, pois não se pode determinar com antecipação o sexo de uma pessoa que fará parte da amostra.

$$X_1 : \Omega_1 = \{\text{"A Estudante é do Sexo Feminino"}; \text{"O estudante é do Sexo Masculino"}\} \rightarrow \text{Imagem}(X) = \{1, 0\}$$

O experimento de se perguntar a um estudante o tipo de segundo grau por ele cursado faz nascer a variável aleatória:

$$X_{2A} = \text{Tipo de 2º Grau Cursado pelos Estudantes}$$

que associa a cada elemento do espaço amostral Ω_2 um número de 1 até 6, de acordo com a regra:

Cursou o Segundo Grau:	Valor de X_{1A}
Majoritariamente em Escola Pública Estadual	1
Totalmente em Escola Pública Estadual	2
Majoritariamente em Escola Pública CEFET	3
Totalmente em Escola Pública CEFET	4

Majoritariamente em Escola Particular	5
Totalmente em Escola Particular	6

Se o pesquisador quiser classificar os estudantes em uma dentre as duas categorias:

$$\left\{ \begin{array}{l} \text{Fez o 2º Grau Total ou Majoritariamente em Escola Pública} \\ \text{Fez o 2º Grau Total ou Majoritariamente em Escola Particular} \end{array} \right\}$$

deverá lançar mão da variável aleatória:

$$X_{2B} = 2^\circ \text{ Grau Majoritariamente Cursado pelos Estudantes.}$$

a qual atuará de acordo com a seguinte regra de associação:

Cursou o Segundo Grau:	Valor de X_{1A}
Majoritariamente em Escola Pública Estadual	1
Totalmente em Escola Pública Estadual	1
Majoritariamente em Escola Pública CEFET	1
Totalmente em Escola Pública CEFET	1
Majoritariamente em Escola Particular	2
Totalmente em Escola Particular	2

As duas variáveis aleatórias X_{2A} e X_{2B} terão por domínio o mesmo espaço amostral Ω_2 mas suas imagens serão diferentes conjuntos:

$$\text{Imagem}(X_{2A}) = \{1, 2, 3, 4, 5, 6\}$$

e:

$$\text{Imagem}(X_{2B}) = \{1, 2\}$$

Em linguagem matemática:

$$X_{2A} : \Omega_2 \rightarrow \{1, 2\}$$

e:

$$X_{2B} : \Omega_2 \rightarrow \{1, 2, 3, 4, 5, 6\}$$

1-11. Variáveis Qualitativas Nominais.

Nos dois exemplos examinados, a imagem da variável aleatória não é um conjunto numérico em seu sentido quantitativo, mas sim um conjunto finito cujos elementos numéricos atuam como **rótulos** de qualidades ou atributos de certo estudante. No primeiro exemplo, o sexo determina a existência de **2** qualidades ou atributos possíveis. No segundo, o tipo de segundo grau cursado dá origem a duas classificações, uma com **6** qualidades ou atributos e outra com apenas **2** atributos ou qualidades.

Variáveis aleatórias como as duas por acima examinadas não medem a realidade objetiva no sentido estrito da palavra, mas **classificam** os fatos materiais entre determinadas **categorias** previamente estabelecidas.

Chamamos tais variáveis de **qualitativas**, em oposição às variáveis quantitativas que serão estudadas mais à frente, uma vez que expressam diferentes **qualidades** pertinentes aos elementos da amostra ou da população. Também as chamamos de **nominais**, pois a

ação de tais variáveis é dar diferentes **nomes** a elementos classificados em diferentes categorias. No caso, os nomes utilizados foram os rótulos numéricos, **1 e 0** ou **1, 2, 3, 4, 5 e 6** ou ainda **1 e 2**. Os números utilizados poderiam ser quaisquer outros e não haveria a menor diferença. Por exemplo, as imagens das variáveis X_1 e X_{2A} poderiam perfeitamente ser iguais ao conjunto $\{0, 1\}$ sem que houvesse qualquer alteração no sentido geral de tais variáveis aleatórias.

Os nomes ou rótulos de uma variável qualitativa nominal devem expressar **categorias exaustivas e excludentes**. O conjunto de categorias deve ser tal que todo elemento da amostra pode ser classificado em alguma delas. E, além disso, cada elemento da amostra só pode ser classificado em uma única categoria.

Estamos dizendo elementos da amostra, mas tal regra valeria igualmente para todos os elementos da população, se existisse condições de classificar todos eles. Um elemento ou pertence ou não pertence a determinada categoria, não há meio termo. Dessa forma, no tocante a sua classificação, os elementos ou são iguais entre si ou são totalmente diferentes. Não existe uma terceira situação, como não existe uma graduação entre as distintas categorias ou qualidades.

O conjunto imagem de uma variável aleatória qualitativa nominal não permite qualquer hierarquização ou ordenamento entre as diversas categorias. Não há sentido algum em se dizer que o sexo masculino é “superior” ao feminino, como não há sentido em se dizer o contrário. Também não há sentido em se dizer que cursar o segundo grau totalmente em escola pública seria “superior” a tê-lo feito em escola particular.

Tanto faz que seja $\{0, 1\}$ ou $\{1, 0\}$ o conjunto de números utilizados como rótulo das categorias “sexo masculino” e “sexo feminino”. Da mesma forma, não é possível hierarquizar o tipo de segundo grau cursado por um estudante universitário e os rótulos $\{0, 1, 2, 3, 4, 5, 6\}$ não refletem qualquer ordem especial.

Visto esse primeiro tipo de variável aleatória, prossigamos com a definição de variáveis aleatórias a partir dos experimentos aleatórios e espaços amostrais estudados.

1-12. Variáveis Qualitativas Ordinais.

A seguinte variável aleatória se origina no espaço amostral do experimento aleatório definido pela pergunta que visa verificar o nível sócio econômico do entrevistado no questionário do “Critério Brasil – IBOPE”:

X_3 = Nível de Instrução do “Chefe de Família” do Entrevistado

a qual leva o referido espaço amostral, seu domínio, até a imagem $\{1, 2, 3, 4, 5\}$ de acordo com a seguinte regra de associação:

O “Chefe de Família”:	Valor de X_3
É Analfabeto ou cursou incompletamente o Primário	1
Completo o Primário ou cursou incompletamente o Ginásial	2
Completo o Ginásial ou cursou incompletamente o Colegial	3
Completo o Colegial ou cursou incompletamente o Superior	4
Completo o curso Superior	5

de forma que podemos escrever:

$$X_3 : \Omega_3 \rightarrow \text{Imagem}(X_3) = \{1, 2, 3, 4, 5\}$$

Essa variável é também **qualitativa**, no sentido que as categorias de sua imagem expressem diferentes qualidades possíveis dos elementos que serão submetidos ao experimento aleatório. No entanto, além de dar nome a cada categoria, essa variável atribui certa hierarquia ao conjunto de categorias elaboradas para classificar a realidade.

Para se ter completado o primário ou cursado incompletamente o ginásial é preciso fazer alguma coisa a mais do que ter cursado incompletamente o primário ou ser analfabeto. O nível $X = 2$, dessa forma, é em alguma medida “superior” ao nível $X = 1$, pois pertencer ao nível $X = 2$ significa ter satisfeito ao nível $X = 1$ e realizado algo a mais. E da mesma forma todos os demais níveis. Para atingir o nível $X = 5$, curso superior completo, é preciso atingir o nível $X = 4$ e fazer alguma coisa a mais do que ter cursado incompletamente o superior ou ter completado o colegial. Para atingir o nível $X = 4$ necessita-se ter feito algo a mais do que passado pelo nível $X = 3$ o qual por sua vez requer algo a mais do que já ter completado o nível $X = 2$ e este requer algo a mais do que o nível $X = 1$.

Há um encadeamento lógico dos níveis, o que resulta em uma hierarquização, um sentido definido de ordem entre as diferentes categorias e, devido a isso, chamamos variáveis como a que estamos estudando de **ordinais**.

Os números utilizados no conjunto imagem de uma variável ordinal devem ser escolhidos de forma tal a representar certa ordem. Uma vez isso satisfeito, qualquer escolha de números é adequada. Assim, o conjunto imagem da variável que estamos estudando poderia ser também:

$$\text{Imagem}(X_3) = \{0, 1, 2, 3, 4\}$$

Para colocar alguma coisa em ordem, necessitamos em primeiro lugar classificar tal coisa segundo nomes diferentes. Logo, **toda variável aleatória ordinal é também nominal**, não sendo a recíproca verdadeira.

1-13. Variáveis Quantitativas Intervalares.

Se estabelecem uma ordem entre os elementos da realidade que classificam, as categorias das variáveis ordinais não estabelecem **distâncias** entre tais elementos. Não há uma diferença comum entre os diversos níveis de instrução que possa servir de unidade de medida para se estabelecer distâncias entre os níveis.

Examinemos agora o experimento aleatório definido pelo questionário fechado acerca das **12 qualidades** que o eleitor julga se certo político possui ou não possui. Podemos englobar todas as perguntas e definir uma única variável aleatória:

$$X_4 = \text{Número de Qualidades que um Político Possui na Opinião do Eleitor.}$$

a qual irá associar a cada elemento do espaço amostral Ω_3 um número de sua imagem:

$$\text{Imagem}(X_4) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

ou seja:

$$X_4: \Omega_3 \rightarrow \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

No quadro conceptual que preside a definição dessa variável, o importante para o estudo do objeto de pesquisa é o **número de qualidades** que o eleitor julga o político possuir, e não **cada uma das qualidades em si**. Se adotamos uma variável assim definida é porque faz parte dos objetivos de pesquisa estudar a quantidade de qualidades de um político, ao contrário de se estudar as qualidades individualmente.

Temos uma variável aleatória **quantitativa**, em oposição às variáveis qualitativas, no sentido de que os números pertencentes à sua imagem $\{0, 1, 2, \dots, 12\}$ representam diferentes quantidades de um mesmo atributo, no caso, quantidade de qualidades que o político possui. O atributo é o mesmo: qualidades de um político, o que muda é a quantidade desse atributo que o eleitor considera o político possuir.

Em cada pergunta do questionário, o eleitor só pode responder sim ou não. Assim, um político só poderá possuir um número **inteiro** de qualidades, oriundo da contagem de quantos **sim** foram respondidos às **12** perguntas. Não poderá possuir **2,5** qualidades ou **3,14** qualidades. Contar significa poder adicionar mais **1**, ou seja, somar uma quantidade sempre fixa e igual, a qual denominamos unidade de medida. Sendo oriunda diretamente de uma contagem, a variável X_4 só pode variar de **1** em **1**. Temos uma variável aleatória **discreta**, em oposição às variáveis contínuas, as quais veremos adiante.

Dentro de um quadro conceptual onde cada qualidade em si não é importante mas sim a quantidade delas, as qualidades de um político podem todas ser tomadas como sendo equivalentes entre si. A partir disso, podemos dizer que a diferença entre ter **8** qualidades, quaisquer que sejam elas, e ter **4** qualidades é a mesma diferença que entre ter **11** qualidades e ter **7** qualidades, e que essa diferença é o dobro da diferença entre ter **3** qualidades e ter **1** qualidade.

Temos uma variável **intervalar**, uma vez que sua imagem possui um **intervalo constante** de variação entre um ponto e outro. Os números $\{0, 1, 2, \dots, 11, 12\}$ não apenas dão ordem às diferentes categorias correspondentes à imagem da variável, como também permitem que se estabeleçam **diferenças** entre duas categorias da imagem.

Para ser intervalar, uma variável deve antes de mais nada poder ordenar as categorias que correspondem a sua imagem. Dessa forma, **toda variável intervalar é também ordinal e em consequência, é também nominal**. De fato, quem conta alguma coisa tem que distinguir o que está contando, ou seja, tem que nomear caso a caso e as distintas quantidades podem ser dispostas em ordem crescente.

O que distingue uma variável aleatória intervalar de uma ordinal é a possibilidade de **contar** ou quantificar certa qualidade. Essa possibilidade é decorrente da existência de um intervalo comum entre as diversas categorias pertencentes a **imagem** da variável, a qual atua como uma **escala de medida** da qualidade que se quer quantificar.

A existência desse intervalo comum entre os valores da variável aleatória permite que se ordene as diferenças que podem existir entre dois valores quaisquer. Assim, a diferença entre um político que possui dez qualidades e outro que só possua duas (na opinião de

certo eleitor) é maior que a diferença entre um político que possui nove qualidades e outro que possui sete.

1-14. O Zero é uma Convenção na Escala Intervalar.

A definição de uma variável aleatória deve ser resultado da elaboração de um quadro conceptual que situe o objeto de pesquisa e esclareça como será realizada a análise estatística para a obtenção de respostas aos objetivos de pesquisa. Dentro do quadro conceptual que esboçamos para a análise de um político, há uma sutil diferença entre as duas variáveis aleatórias seguintes:

X_4 = Número de Qualidades que o Político Possui na Opinião do Eleitor.

X_{4A} = Número de Respostas “Sim” que o Eleitor Marcou no Questionário.

Embora a imagem das duas variáveis seja a mesma,

$$\text{Imagem}(X_4) = \text{Imagem}(X_{4A}) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\},$$

existe a diferença seguinte: Em função da definição da primeira variável ter sido realizada dentro de um quadro conceptual, **o valor zero de sua imagem é uma convenção.**

De fato, se alguém atribuiu **zero** respostas “sim” ao político, isso não significa necessariamente que essa pessoa considere ter ele **zero qualidades**. Mesmo quem tenha respondido **não** às **12** perguntas certamente concordará que o político possui outras qualidades além das **12** listadas no questionário. O **zero** da escala de medida dessa variável, em consequência não significa a **ausência de quantidade**, não significa um **conjunto vazio**.

Um eleitor que respondeu ter certo político **10** qualidades não está lhe atribuindo o dobro de qualidades que a outro, possuidor de **5** qualidade. Não havendo um zero verdadeiro na escala de medida das quantidades, não é possível estabelecer **relações de razão** entre tais quantidades.

As variáveis intervalares permitem ordenar categorias e permitem também estabelecer diferenças entre categorias distintas, diferenças essas que podem ser também ordenadas. Mas para aí e não permitem estabelecer proporções entre duas categorias distintas.

1-15. Variáveis Quantitativas Proporcionais.

O espaço amostral do experimento de se tomar um ônibus previamente sorteado e contar o número de passageiros resulta na seguinte variável aleatória:

X_7 = Número de Passageiros em um Ônibus

a qual associa a cada elemento do espaço amostral Ω_{6B} um número da seguinte imagem:

$$\text{Imagem}(X_7) = \{0, 1, 2, 3, 4, 5, \dots\}$$

Temos novamente uma variável aleatória **quantitativa e discreta**, mas agora o zero de sua escala de medida possui o significado de um conjunto vazio, ou seja, ausência de

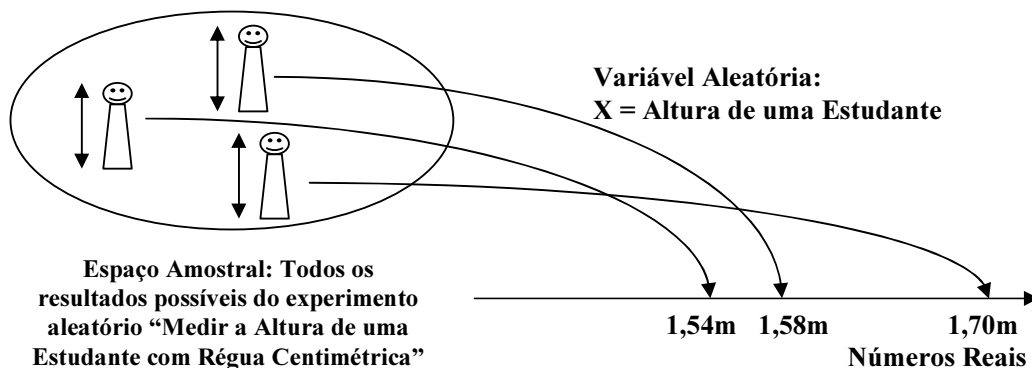
passageiros em um ônibus. Em conseqüência, podemos dizer que certo ônibus está carregando o dobro ou o triplo de passageiros que outro. Podemos tomar proporções com os valores dessa variável e comparar tais proporções com as realizadas para outras variáveis. Se um ônibus carrega cinco vezes mais passageiros que outro essa proporção é a mesma que existe entre o volume de cinco garrafas e de uma garrafa de cerveja.

Temos então uma variável **proporcional**, que é capaz de estabelecer proporções entre diferentes aspectos da realidade objetiva. Tais variáveis podem ser discretas, como o número de passageiros em um ônibus ou contínuas, como no exemplo a seguir.

O experimento aleatório de se medir a altura de uma estudante com uma régua centimétrica origina a variável aleatória:

$$X_7 = \text{Altura de uma Estudante.}$$

A figura abaixo ilustra o mecanismo de associação unívoca dessa variável aleatória:



Se abstrairmos a precisão do instrumento de medida, a altura de uma estudante pode ser considerada uma variável aleatória **contínua**, podendo assumir qualquer valor dentro de determinado intervalo. Medindo com uma régua de precisão **centimétrica** (a distância entre dois traços da régua é **1cm**), teríamos, por exemplo, **1,60m**; **1,61m**; **1,62m** etc. Mas se medíssemos com uma régua de precisão **milimétrica** (a distância entre dois traços da régua é **1mm**), teríamos **1,600m**; **1,601m**; **1,602m** etc.

Idealizando o processo, a precisão do instrumento poderia ser continuamente aumentada, resultando no conceito de **números reais**, números tais que entre dois deles, por mais próximos que estejam, sempre existe um terceiro, maior que o primeiro e menor que o segundo. Chamamos de **variáveis aleatória contínuas** aquelas que podem assumir valores em um **intervalo contínuo**, ou seja, que têm por imagem um **intervalo real**.

Todas as variáveis aleatórias ditas contínuas são, na verdade, discretas. Vejamos o porquê dessa aparente contradição. Como todos os instrumentos de medida tem precisão limitada, somente medidas discretas são possíveis, visto que o processo de medição conta quantas vezes certa dimensão é maior que a escala do instrumento. Não é possível, por exemplo, com uma régua metálica, medir a altura de pessoas com precisão de **0,1mm** ou seja, **0,0001m**. As balanças comerciais mostram esse fato claramente, pois os resultados de medida são apresentados em "displays" eletrônicos com certo número

fixo de dígitos. Para facilitar as manipulações matemáticas, no entanto, consideramos que uma variável é contínua se pudermos idealizar sua precisão de medida como infinita.

Toda variável proporcional é também intervalar, sendo consequentemente ordinal e nominal. A recíproca, como é claro, não é verdadeira.

1-16. Escalas de Medida.

Vimos que a Estatística inicia seu trabalho com os números obtidos de uma amostra sendo para isso necessário classificar ou medir os elementos da amostra através de algum critério, **mensurar a realidade material ou humana, transformando-a em números**, os quais serão posteriormente analisados.

A imagem de uma variável é uma **escala de medida**, no sentido amplo da palavra, do fenômeno aleatório descrito pelo espaço amostral. As variáveis aleatórias mensuram, de alguma maneira, a realidade objetiva. Como vimos, existem **quatro níveis possíveis de mensuração a realidade** e a cada nível corresponde um tipo de escala de medida.

No nível mais elementar de mensuração, podemos apenas distinguir se dois elementos são iguais ou diferentes entre si, em termos de algum conjunto de características de interesse, de forma a **nomear** os elementos estudados como pertencentes a diferentes categorias. Consideramos todos os elementos da cada categoria como iguais entre si, em relação às características que definem a categoria. Essa escala de medida é chamada de **nominal** e possui as seguintes propriedades que permitem a classificação dos elementos da realidade em estudo:

- (1.) $a = b$ ou $a \neq b$
- (2.) Se $a = b$ então $b = a$
- (3.) Se $a = b$ e $b = c$ então $a = c$

Em um segundo nível, podemos não apenas distinguir entre dois elementos iguais ou diferentes mas também colocá-los em ordem segundo algum critério. A escala de medida é chamada de **ordinal** e possui as propriedades que se seguem, além das propriedades da escala nominal:

- (4.) Se $a < b$, então $b > a$
- (5.) Se $a < b$ e $b < c$ então $a < c$

Em um terceiro nível de complexidade, além de distinguir e ordenar dois elementos podemos também estabelecer diferenças entre dois elementos quaisquer, diferenças essas possíveis de serem ordenadas. Existe um intervalo mínimo que pode ser tomado como unidade de medida entre elementos e a escala é chamada de **intervalar**. Podemos realizar a operação de **adicionar** dois elementos para obter um terceiro, correspondendo às propriedades seguintes:

- (6.) $a + b = b + a$
- (7.) Se $a = b$ e $c = d$ então $a + c = b + d$
- (8.) $(a + b) + c = a + (b + c)$
- (9.) $\exists 0$ tal que $a + 0 = a$

Note o leitor que a propriedade **9** define o zero como um ponto de partida para a adição de valores, o que corresponde a noção de origem por convenção.

No quarto e último nível de complexidade, além de distinguir, ordenar e estabelecer diferenças ordenáveis entre dois elementos, podemos também estabelecer proporções entre elementos distintos e a escala é chamada de proporcional. Podemos realizar a operação de **multiplicação** e a propriedade correspondente, embora não pareça a primeira vista, dá sentido absoluto à existência do elemento neutro na escala intervalar. Com efeito, na presença dessa propriedade, o zero deixa de ser uma convenção:

$$(10.) \exists 1 \text{ tal que se } a \cdot b = 1 \text{ então } a = 1/b$$

Estabelecemos, dessa forma, uma **escala ordinal** para classificar os diferentes tipos de variáveis aleatórias:

X = Tipo de Uma Variável Aleatória Qualquer.

Variável	X
Nominais;	1
Ordinais;	2
Intervalares;	3
Proporcionais.	4

havendo sentido na expressão: $1 < 2 < 3 < 4$

1-17. Qualidade e Quantidade.

Uma **variável aleatória** é **qualitativa** quando seus valores representam **qualidades** ou **atributos** que se queira estudar em uma pesquisa. Uma variável qualitativa é sempre **discreta**, visto que entre uma qualidade e outro não são admitidas gradações. É branco ou preto, é pau ou pedra, é masculino ou feminino.

Uma variável aleatória é **quantitativa**, quando expressa seus valores em termos **numéricos**. Será **discreta** se originar-se em alguma forma de **contagem**, sendo sua imagem um conjunto **discreto**, ou seja, que contenha uma quantidade finita ou enumerável de valores numéricos. Será **contínua** se originar-se, em alguma forma de medida, sendo sua imagem um intervalo **real**. Enquanto as variáveis ordinais e nominais são sempre **qualitativas**, as variáveis intervalares e proporcionais são sempre **quantitativas**.

As variáveis qualitativas, classificando os elementos de uma amostra ou de uma população entre atributos ou qualidades diferentes, sempre podem dar origem a variáveis aleatórias quantitativas, as quais representam a **frequência** ou o **número de elementos classificados em cada atributo ou qualidade**.

No exemplo da variável aleatória qualitativa que classificava um estudante sorteado para fazer parte de uma amostra, temos duas variáveis aleatórias discretas:

f_M = “Número de Estudantes de Sexo Masculino da Amostra”

f_F = “Número de Estudantes de Sexo Feminino da Amostra”

de onde se deduz que:

$$f_M + f_F = \text{Tamanho da Amostra} = n$$

Além disso, podemos ter também duas outras variáveis aleatórias:

fr_M = “Frequência Percentual de Estudantes de Sexo Masculino na Amostra”

fr_F = “Frequência Percentual de Estudantes de Sexo Feminino na Amostra”

que podem ser obtidas através das expressões seguintes:

$$fr_M = \frac{f_M}{n} 100\% \quad \text{e:} \quad fr_F = \frac{f_F}{n} 100\%$$

de onde:

$$fr_M + fr_F = 100\% = 1$$

Essas duas variáveis parecem ser contínuas, já que seus valores são obtidos através de contas matemáticas e possuem casas depois da vírgula. No entanto, embora não variem de 1 em 1, elas se originam de um processo de contagem e são discretas. De fato, a imagem de tais variáveis aleatórias é sempre um conjunto **discreto** de valores **racionais** ou seja, do conjunto que pode ser representado na forma da razão p/q , sendo p e q inteiros.

1-18. Valores Amostrais e Populacionais de uma Variável Aleatória.

Seja a variável aleatória contínua:

X = Altura de uma Estudante.

Existe na UFBA certo número N_F de estudantes femininas. O tamanho exato N_F dessa população pode ser obtido na Secretaria Geral de Cursos. Dessa forma, existe o mesmo número de **alturas** de estudantes femininas. Se realizássemos um censo e medíssemos todas as alturas dessas estudantes, tais medidas numéricas se constituiriam nos **valores populacionais** da variável aleatória **X**:

$$X \in \{x_1, x_2, x_3, \dots, x_{N_F}\}$$

Se **31** alunas de uma **amostra** representativa e aleatoriamente coletada dessa população tiverem suas altura medidas, teremos **31** valores de altura feminina, os quais serão **valores amostrais** da variável aleatória **X**. Podemos ordenar tais valores amostrais, obtendo um conjunto chamado de **rol**:

$$X \in \{1,48; 1,52; 1,53; 1,54; 1,55; 1,56; 1,57; 1,57; 1,57; 1,58; 1,58; 1,59; 1,60; 1,60; 1,60; 1,60; 1,60; 1,61; 1,61; 1,62; 1,62; 1,62; 1,62; 1,65; 1,68; 1,69; 1,70; 1,70; 1,75; 1,76; 1,78; 1,78\}$$

Os **valores amostrais** de uma variável aleatória são aqueles provenientes da coleta de dados em uma **amostra**. Já os **valores populacionais** são originados em um **censo**, ou coleta de dados realizada em toda a **população**. Note-se que **mesmo que tal censo não possa ser realizado, nem por isso os valores populacionais deixam de existir**. De fato, as estudantes da UFBA não deixam de ter cada uma sua altura apenas por que ninguém realizou um censo para as medir uma a uma.

Vimos que a variável altura possui um grande número de valores populacionais e um número determinado de valores amostrais. Já a variável

X = “Frequência percentual de estudantes do Sexo Feminino”

possui um único valor populacional, que só pode ser obtido se o sexo de cada elemento da população for determinado. Cada amostra de estudantes, independente de seu tamanho, possui também um único valor amostral dessa variável, o qual pode ser obtido determinando o sexo de todos os estudantes da amostra.

1-19. A Estatística Descritiva.

A importância da distinção entre valores amostrais e valores populacionais de uma mesma variável aleatória reside em sua utilização no âmbito de uma pesquisa. Os primeiros constituem os dados de partida, que serão descritos matematicamente e analisados para possibilitar a realização de inferências ou previsões acerca dos segundos.

A **estatística descritiva** é a que trabalha com os **valores amostrais** das variáveis aleatórias, aplicando-lhes uma série de fórmulas que calculam as chamadas “**estatísticas descritivas da amostra**”

No caso do objeto de estudo ser a altura das estudantes femininas, por exemplo, uma série de contas matemáticas é realizada com os valores amostrais dessa variável aleatória, obtendo-se a altura média, a moda da altura, a altura da estudante mediana, as alturas das estudantes situadas nos quartis, o desvio quartílico da altura, o desvio padrão da altura, o desvio padrão da altura média e a distribuição de frequência das alturas.

Por exemplo, a **altura média** das estudantes da referida amostra pode ser calculada como $\bar{x} = 1,62\text{m}$ e o **desvio padrão das alturas individuais** resulta em $s = 7,7\text{ cm}$, o que acarreta um **desvio padrão da altura média** igual a $s_n = 1,4\text{ cm}$.

A estatística descritiva fornece as fórmulas matemáticas para se calcular as **estatísticas descritivas amostrais**, que são as já citadas média, etc, bem como fornece a **interpretação** em português de seus significados. É de especial importância descrever minuciosamente a amostra em termos do significado conceptual das estatísticas descritivas e tal só pode ser realizado em português.

De fato, qual o sentido em se calcular um monte de coisas cujo significado desconhecemos? A resposta é óbvia: só há sentido em calcular alguma coisa se for para **interpretar** o valor obtido. E esse é o campo de atuação da **Estatística Descritiva**.

1-20. A Estatística Indutiva.

Se a estatística descritiva descreve o comportamento de **31** valores amostrais da variável **X = Altura de uma Estudante**, o que dizer de todos os valores populacionais dessa variável aleatória, ou seja, das alturas de todas as estudantes da UFBA? Por exemplo, o que poderia ser afirmado sobre a **altura média μ de todas das estudantes da UFBA**?

A **estatística indutiva** permite estabelecer conclusões populacionais a partir das estatísticas descritivas amostrais. No exemplo que estamos discutindo, além da amostra de alturas femininas ser representativa da população, devemos considerar que a variável aleatória

$$X = \text{Altura de uma Estudante.}$$

possui um comportamento **normal**, o que permite os cálculos através de consultas a tabelas de probabilidades como a do *Anexo 1* do *Livro 6 – O Modelo Normal de Probabilidades*.

Podemos afirmar inicialmente que o valor da altura média amostral $\bar{x} = 1,62\text{m}$ seria a melhor **estimativa pontual** possível para a altura média populacional de todas as estudantes da UFBA. Escrevemos em linguagem matemática:

$$\text{Média}(X) = \mu \approx \bar{x} = 1,62\text{m.}$$

e lemos em português: a média populacional das alturas **é estimada** pelo valor da média amostral, a qual é **igual** a **1,62m**. Note-se a diferença entre o símbolo $\approx$ (é **estimado**) e o símbolo $=$ (**igual**). Não podemos dizer que a média populacional é **igual** a **1,62m** pelo simples fato de que uma **outra amostra** poderia resultar em uma **média amostral diferente** e, portanto, uma **estimativa pontual diferente** para a **média populacional**.

Como foi calculado o desvio padrão da altura média $s_n = 1,4\text{ cm}$, a Estatística permitiria induzir que uma **margem de erro** de **3cm** em torno da média amostral **1,62m** teria um **nível de confiança** igual a **97%**. Explicando melhor: um intervalo que se iniciasse em $1,62\text{m} - 3\text{cm} = 1,59\text{m}$ e terminasse em $1,62\text{m} + 3\text{cm} = 1,65\text{m}$ iria conter a altura média populacional μ das estudantes com probabilidade igual a **97%**.

Em linguagem matemática, para a **altura média populacional** μ temos uma **estimativa por intervalo**, o qual é construído através de uma **margem de erro** de $e_0 = 3\text{cm}$ em torno da média amostral $\bar{x} = 1,62\text{m}$ e possui um **nível de confiança** de $1 - \alpha = 97\%$:

$$P(\bar{x} - e_0 \leq \mu \leq \bar{x} + e_0) = 97\%$$

Interpretando essa expressão, temos **confiança** de que, se coletarmos muitas amostras aleatórias de tamanho $n = 31$, em **97%** das vezes um intervalo formado por uma **margem de erro** de $e_0 = 3\text{cm}$ para mais ou para menos em torno da altura média amostral $\bar{x}$ irá conter a altura média populacional feminina μ .

Se aumentássemos o nível de confiança, a precisão do resultado iria diminuir, visto que o intervalo em torno da média amostral seria maior. E se diminuíssemos a margem de erro para aumentar a precisão, perderíamos em confiança. Em qualquer caso, só teríamos **certeza** do valor exato da altura média populacional feminina se realizássemos um **censo**, medindo todas as estudantes da UFBA e calculando a média de todas elas.

1-21. A Probabilidade É uma Previsão do Futuro.

Alguna leitora deve estar se perguntando: “–E eu, que tenho **1,57m**?” Note a prezada leitora que até agora estávamos falando de **altura média de todas as estudantes** e não

de alturas individuais de cada uma. Passemos então a ilustrar o que a estatística permite induzir acerca das alturas individuais das estudantes da população pesquisada.

Se a leitora possui **1,57m** de altura, poderíamos calcular a probabilidade de existirem estudantes iguais ou mais baixinhas que ela. Essa probabilidade seria **25,8%**. Dito de outra forma, com base na amostra coletada podemos inferir que a prezada amiga é mais alta que aproximadamente $\frac{1}{4}$ da população feminina. Em linguagem matemática:

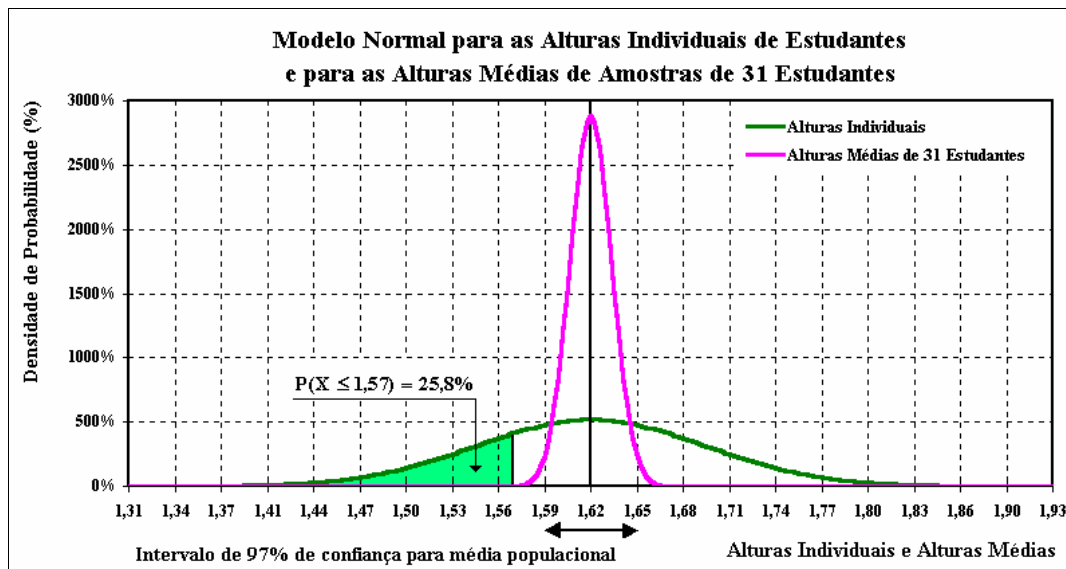
$$P(X \leq 1,57) = 25,8\%$$

Essa afirmação não é uma certeza, mas um prognóstico, uma previsão do futuro. Não se conhece todos os valores populacionais da altura, mas podemos prever que se um censo for realizado, **25,8%** da população de estudantes será mais baixinha que a prezada leitora com seus **1,57m**.

Outra espécie de previsão que pode ser feita: se sortearmos uma estudante ao acaso e compararmos sua altura com **1,57m**, a longo prazo, em média, em **25,8%** das vezes tal estudante sorteada será mais baixinha que a leitora.

O gráfico abaixo ilustra a situação. A curva a curva lilás e pontiaguda representa o modelo normal da distribuição das alturas médias de amostras de tamanho **31**. Foi através de valores originados dessa curva lilás que calculamos a margem de erro e o nível de confiança da estimação da **média populacional** das alturas femininas.

A curva verde e achatada, por seu lado, representa o modelo normal da distribuição populacional das alturas individuais. Foi através dela que calculamos a probabilidade de existirem estudantes mais baixinhas que **1,57m**.



Se a prezada leitora não possui **1,57m**, verifique o exato ponto onde sua altura se encontra! Basta entrar com o valor no eixo horizontal do gráfico e observar sua posição em relação à curva verde: perto da média, na cauda direita ou na cauda esquerda? Traçando a área sob a curva verde à esquerda do ponto onde está sua altura, a leitora terá uma

idéia da porcentagem de estudantes da UFBA menores que ela. Traçando a área à direita, observará a porcentagem de estudantes que lhe são mais altas. Faça isso!

Veja a força da Estatística! Através de uma amostra de apenas **31** alturas de estudantes femininas, podemos traçar conclusões para as alturas todas as estudantes da UFBA! No entanto, para isso dar certo, a amostra deve ser **representativa** da população, deve ser **imparcial** e possuir **tamanho** suficiente. Esse é o tema que estudaremos a seguir.

1-22. A Teoria da Amostragem.

A principal questão que se deve responder antes de se proceder aos cálculos da estatística indutiva é: serão os valores amostrais da variável aleatória em estudo capazes de fornecer estimativas confiáveis de seus correspondentes valores populacionais? A **Teoria da Amostragem** é a parte da Estatística que estuda os métodos utilizados para a obtenção de amostras que possibilitem a indução dos resultados para a população no tocante a certa variável aleatória em estudo.

Uma amostra, sendo parte do todo populacional, nunca será uma miniatura perfeita da população, sempre haverá um certo desvio entre as estatísticas amostrais e os parâmetros populacionais. O importante na teoria da amostragem é que a amostra seja **representativa** da população, no sentido de espelhar o mais fielmente possível o comportamento populacional.

O principal condicionante da representatividade de uma amostra face certa população, quando se pretende estudar o comportamento de determinada variável aleatória é o **perfil** da amostra em relação aos fatores que influenciam o comportamento dessa variável. Esses fatores, em relação aos quais é importante a igualdade de perfil entre população e amostra, são chamados de **variáveis de controle** de uma pesquisa. Quanto mais próximo o perfil amostral das variáveis de controle é do perfil populacional, maior a representatividade da amostra em relação à variável aleatória em estudo.

O **processo de escolha** dos elementos da amostra determina seu caráter **imparcial** ou **parcial**. Já mencionamos a amostragem **aleatória**. Em relação ao processo de escolha dos elementos da amostra, é a que melhor garante a **imparcialidade**. Cada elemento da amostra é obtido através um sorteio onde todos têm igual probabilidade de serem escolhidos e todas as possíveis amostras têm igual probabilidade de serem extraídas da população. As vezes, a amostra aleatória tem que ser subdivida, de forma a representar mais adequadamente a **estratificação** de diversos setores da população.

No processo de amostragem chamado de **à esmo**, o pesquisador escolhe a amostra sem qualquer regra aparente, com exceção àquelas utilizadas por seu subconsciente.

Existe também o processo por **preenchimento de cotas**, que é derivado do processo à esmo, mas que atende a critérios tais que a amostra resulta possuidora de um perfil o mais próximo possível ao perfil populacional, no tocante aos fatores que mais influem no comportamento da variável aleatória estudada. Esse processo procura compensar a parcialidade do processo a esmo através de um acréscimo planejado na representatividade da amostra.

1-23. Tamanho não é Documento para Amostras...

Finalmente, o **tamanho** da amostra. Para amostras de mesmo perfil e mesmo processo de escolha, quanto maior o tamanho, menor a margem de erro dos resultados e maior a confiança que se pode depositar em tal margem de erro. O tamanho de uma amostra deve ser dimensionado em função da dispersão existente nos dados amostrais da variável que se está estudando. Para isso, pode ser necessário a coleta de uma amostra piloto.

Em termos de uma amostra, tamanho não é documento. Muito mais importante é a representatividade da amostra face a população. Se a amostra **não for representativa**, poder-se-á realizar um estudo estatístico acerca dos valores amostrais, mas será **impossível** estender as conclusões desse estudo para os valores populacionais.

Uma lenda bastante arraigada entre os estudantes de ciências humanas conta que para ser representativa uma amostra deve ter **10%** da população. Ora, uma amostra de **10%** da população não é “automaticamente” representativa só porque possui esse tamanho. Se o perfil dessa amostra for totalmente diferente do perfil populacional, naquilo que é importante para o comportamento da variável aleatória em estudo, ela poderá englobar até mais do que **10%** da população, mas simplesmente **não será representativa**.

Estudemos brevemente o exemplo das pesquisas eleitorais. Uma amostra de tal pesquisa possui normalmente tamanho **2000**, representando apenas uma ínfima porcentagem do total de eleitores. Mais exatamente, representa a porcentagem de:

$$\frac{2.000}{100.000.000} 100\% = 0,02\%$$

da população de eleitores brasileiros. No entanto, tal amostra é estratificada, subdivida em amostras regionais onde cada uma representa o perfil populacional da região em termos de gênero, idade, escolaridade, nível sócio econômico e relação de tamanho entre os centros urbanos e rurais. Todos os elementos pesquisados são obtidos por sorteio, a amostra é estratificada e aleatória.

Tomadas todas essas precauções, uma amostra de **0,02%** da população é suficientemente representativa para a pesquisa apresentar resultados tecnicamente confiáveis.

1-24. A Construção de Um Quadro Conceptual.

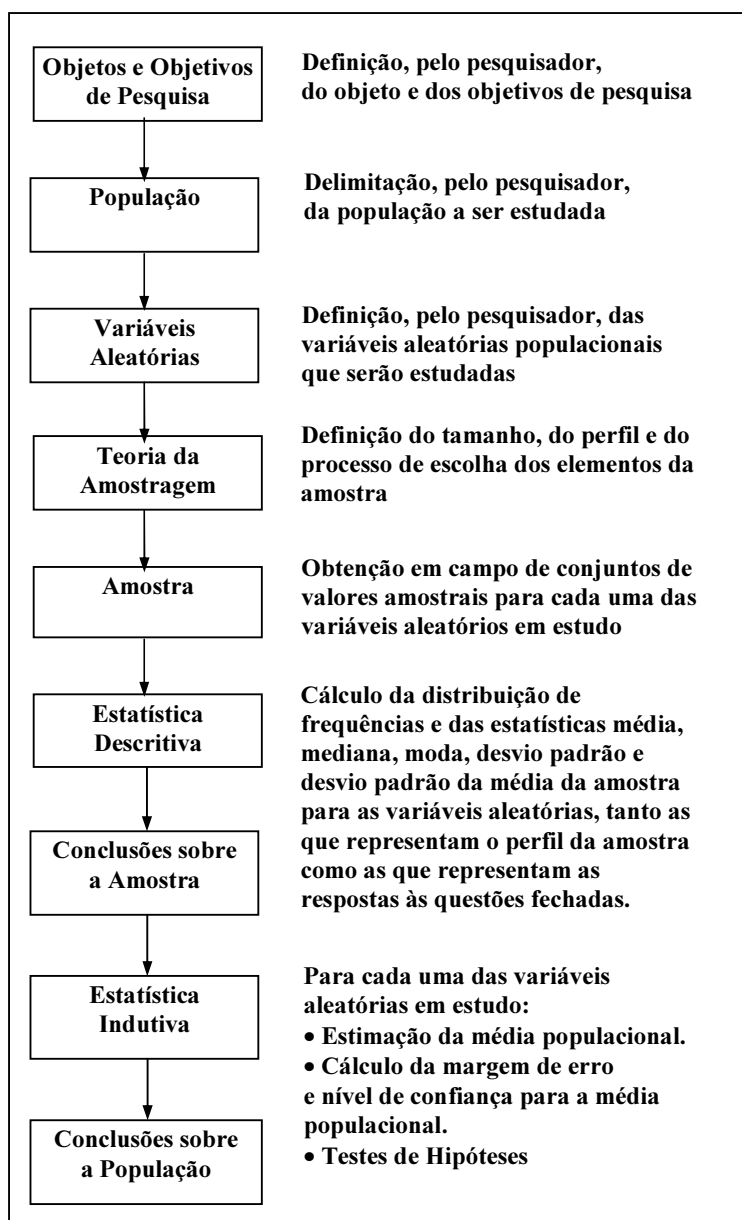
A utilização da Estatística em pesquisas pode ser ilustrada pela “equação”:

$$\boxed{\text{Estatística} = \text{Matemática} + \text{Português}}$$

Os fatos sociais e humanos só podem ser estudados através da construção de conceitos expressos em determinada língua. Utilizaremos o bom e velho Português (!) As palavras servem para a construção de um **quadro conceptual** que permita análise do objeto e dos objetivos de estudo. De acordo com esse quadro conceptual, o pesquisador deverá estabelecer uma série de variáveis aleatórias, definidas em Português, que devem ser adequadas a obtenção de respostas aos seus objetivos de estudo.

Obtendo no campo os dados de uma amostra, o pesquisador deverá realizar os cálculos da estatística descritiva, que lhe trarão conclusões acerca desta, bem como os cálculos da estatística indutiva, que lhe mostrarão até que ponto tais conclusões amostrais são válidas também para a população pesquisada. As conclusões estatísticas necessitarão de palavras em Português para se transformarem em conclusões acerca do objeto de estudo, no âmbito da população delimitada pelo pesquisador e de acordo com o quadro conceitual anteriormente definido.

O diagrama seguinte ilustra o processo de pesquisa estatística, com suas interconexões **matemática x português**.



A importância da construção de um quadro conceitual está em que o conhecimento teórico do pesquisador sobre o objeto de estudo e seu método filosófico de obtenção do conhecimento não pode ser substituído pela análise estatística.

Ao contrário, tal conhecimento deve servir para enquadrar a análise estatística dentro de uma caracterização mais abrangente da realidade, de forma a que esta esclareça situações e reforce ou negue argumentações sobre as relações de causalidade intrínsecas ao objeto de estudo.

O perigo para o pesquisador é a substituição de seu método específico de obtenção de conhecimento por uma caricatura do método estatístico.

Quando isso ocorre em pesquisas ditas “quantitativas”, as variáveis aleatórias são impostas como sendo “os” conceitos em estudo e procura-se estabelecer conclusões e rela-

ções estatísticas entre tais variáveis como sendo “as” conclusões e “as” relações existentes acerca dos conceitos.

1-25. O Método Estatístico

Ocorre que as variáveis aleatórias não são “os” conceitos em si, mas representações estatísticas de aspectos desses conceitos. Em consequência, todo processo de substituição de conceitos por variáveis definidas de modo mais ou menos arbitrários retira a abrangência do método de pesquisa e o reduz ao método estrutural-funcionalista.

Justamente para evitar tais perigos é que insistimos tanto em discutir a relação entre **método filosófico**, o **método científico** e o **método estatístico**. Justamente por isso, um dos principais **objetivos** de nosso curso é habilitar os futuros profissionais ao **diálogo com os estatísticos**, necessário para o projeto, realização e análise de pesquisas que envolvam técnicas estatísticas.

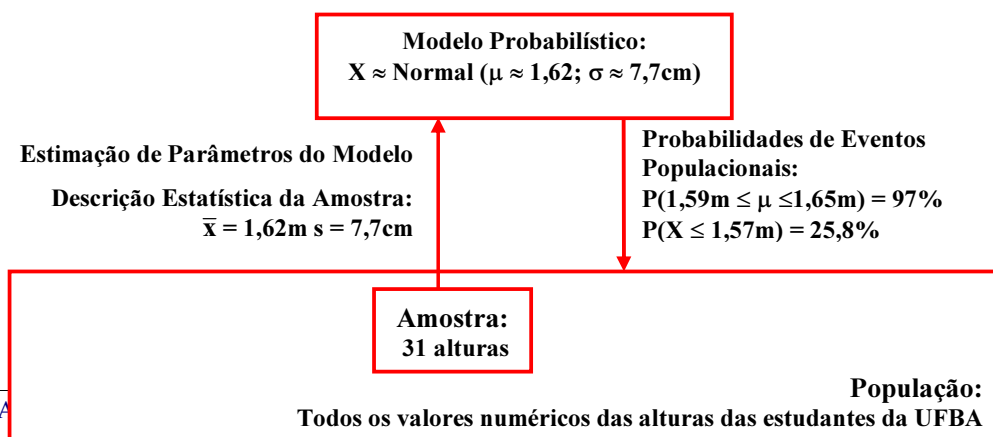
Assim, uma vez definidos os principais conceitos de estatística, voltemos a examinar agora de uma forma mais aprofundada, como o método dessa ciência se enquadra na teoria mais geral de construção do conhecimento humano.

Seja **X** uma variável aleatória que representa uma característica em estudo de certa **população**. Queremos estudar o comportamento dessa característica, mas não podemos, por razões de economia, obter os valores populacionais de **X**. Contornamos essa dificuldade retirando uma **amostra** representativa da população e obtemos um certo número de valores amostrais de **X**. Calculamos, através dos valores amostrais, as **estatísticas descritivas** de **X**.

Como é impossível conhecer todos os valores populacionais de **X**, adotamos um **modelo probabilístico** que represente toda a população, chamado de **distribuição populacional de probabilidades**. Através dos valores que calculamos para as estatísticas descritivas da amostra, **estimamos** os valores dos **parâmetros** desse modelo populacional.

Uma vez estimados os parâmetros populacionais, podemos, através do uso de tabelas, calcular as **probabilidades** de eventos populacionais de interesse, ou seja, de eventos que envolvem os valores populacionais da variável **X**. Tais probabilidades são **previsões da frequências relativas** com que poderão futuramente ocorrer tais eventos populacionais de interesse do pesquisador.

O método estatístico, dessa forma, encerra um ciclo de **indução e dedução**:



Através de uma parte do todo material, ou seja, dos valores de uma amostra da população, **induzimos** uma teoria representativa desse todo, ou seja, um modelo probabilístico populacional. E através dessa idéia ou teoria, **deduzimos** possíveis comportamentos futuros do todo material que estamos estudando, ou seja, calculamos probabilidades de eventos populacionais.

O método estatístico é **materialista**, pois parte de **fatos materiais** (estatísticas descritivas da amostra) para estabelecer teorias (modelos probabilísticos) e não ao contrário. E é **dialético** pois trabalha com o **movimento das idéias**, ao contrário de supor que as idéias são algo estático. Uma vez estabelecidas as idéias que representam a realidade material (modelos probabilísticos), o método estatístico realiza transformações nessas idéias (cálculos de probabilidade) que possibilitam a previsão de futuros comportamentos materiais. E se os comportamentos previstos não ocorrem, os modelos são criticados e substituídos por outros mais adequados, permitindo a evolução da ciência.

1-26. A Importância da Estatística e de seu Estudo.

Por que é importante, a partir dos dados de uma amostra, estabelecer conclusões para certa população? Essa pergunta em si já responde qual é a importância da Estatística. Considere, por exemplo, as seguintes variáveis aleatórias:

X1 = “Porcentagem de Eleitores que Votariam em Certo Candidato”

Qual a importância de podermos inferir a opinião política do conjunto dos eleitores de um estado ou mesmo de todo o Brasil a partir de uma amostra desses eleitores?

X2 = “Renda Média dos Estudantes da UFBA”

Qual a importância de podermos inferir a renda média dos estudantes da UFBA a partir de uma amostra de estudantes?

X3 = “Porcentagem de Desempregados na Região Metropolitana de Salvador”

Qual a importância de podermos inferir o índice de desemprego em Salvador a partir de uma amostra da população da região metropolitana dessa cidade?

Permitindo que se chegue a **conclusões** acerca do **tudo** a partir do estudo de uma **fração** desse todo, a Estatística torna economicamente possível uma série de estudos que não o seriam sem ela. Fundamentalmente nisso é que reside sua importância perante a construção do conhecimento humano, como ferramenta à disposição das diversas ciências.

É importante o estudante compreender perfeitamente a metodologia estatística pois tal compreensão lhe servirá não apenas para a realização de estudos e pesquisas próprios, mas também para desmascarar toda uma série de manipulações que se utilizam da Estatística para impingir à população, como verdadeiras, conclusões falsas.

De fato, a importância de qualquer cidadão compreender o método estatístico está na razão direta da necessidade de explicarmos, com palavras simples e compreensíveis à grande maioria do povo brasileiro, o quão mentirosas e a serviço de quem tais “pesquisas estatísticas” estão.