



Biometria - EDAP

População, Amostra, Variáveis, Dados

(Leitura complementar ao [capítulo 1](#))

População (ou Universo)

- em **Biologia**: é o conjunto de indivíduos de uma certa espécie, que ocupa uma certa área em um determinado intervalo de tempo.
- em **Estatística**: é a totalidade de observações individuais dentro de uma área de amostragem delimitada no espaço e no tempo, sobre as quais serão feitas inferências.

Ou seja, a população é um conjunto de indivíduos ou objetos que apresentam em comum determinadas características definidas para o estudo.

Importante é lembrar que a **tomada de decisões sobre a população**, com base em estudos feitos sobre os dados de uma amostra, constitui a problema central da inferência estatística.

Amostragem

Quando se realiza um estudo **quase nunca** é possível examinar todos os elementos da população em que se está interessado, por motivos financeiros, ou por limite de tempo, ou restrição na locomoção para registro dos dados.

Mas, geralmente, pode-se trabalhar com o que é acessível, ou seja, com uma parte da população, que deve apresentar a maioria das características da população.

Assim, **amostragem** é o processo pelo qual se obtém informação sobre um todo (população), examinando-se apenas uma parte do mesmo (amostra).

Felizmente, a inferência estatística nos dá elementos para generalizar as conclusões obtidas da amostra para toda a população.

Deve-se lembrar que os **erros de coleta e manuseio** de um grande número de dados podem ser maiores do que as imprecisões a que estamos sujeitos quando generalizamos, via inferência, as conclusões de uma amostra bem selecionada. Portanto, é incorreto pensar que seríamos mais precisos, se tivéssemos acesso a todos os elementos da população.

Entretanto, **risco** é a margem de erro motivada pelo fato de se investigar parcialmente (amostra) o universo (população).

Ressalte-se que quanto maior e mais representativa for a amostra, mais forte é o argumento. Se uma amostra for **insuficiente** ou **tendenciosa**, conclusões retiradas a partir dela poderão ser totalmente falsas.

Portanto, em um levantamento amostral deve-se:

- Selecionar a característica a ser pesquisada
- Definir cuidadosamente a população de interesse

Uma das maneiras de se conseguir representatividade é fazer com que o processo de **escolha da amostra** seja, de alguma forma **aleatório**, ou seja, totalmente ao acaso. Portanto, sempre que se puder deve-se planejar o levantamento evitando a parcialidade (**vício**) na seleção.

Tipos de amostragem

Basicamente, existem dois métodos para obtenção da amostra: **probabilístico ou não**.

Nas amostragens **não probabilísticas** os elementos são escolhidos a esmo ou intencionalmente. Esse tipo de amostragem algumas vezes procura usar procedimento aleatório, mas sem um sorteio, ou seja, sem utilizar dispositivos aleatórios confiáveis. Por exemplo: escolher vizinhos ou amigos.

Entretanto, os resultados da amostragem a esmo são, em geral, equivalentes aos de uma amostragem probabilística desde que:

- a população seja homogênea e
- se o amostrador não for inconscientemente influenciado por alguma característica dos elementos da amostra.

Nas amostragens **probabilísticas** as unidades amostrais são escolhidas mediante mecanismos de sorteio. Assim cada elemento da população possui uma certa probabilidade de ser selecionado a qual é, em geral, a mesma para todos os indivíduos. Assim se o tamanho da população for **N**, a probabilidade de um indivíduo estar nela será **1/N**.

De modo geral, para garantir a busca da imparcialidade e representatividade usadas no plano experimental, é preferível escolher os elementos que participarão da amostras aleatoriamente; ou seja usando amostras probabilísticas.

Evidentemente, deve-se evitar o **vício amostral**, que vem da tendenciosidade das informações. O vício ocorre quando existe uma tendência na seleção das unidades amostrais, favorecendo uma dada característica particular.

É preciso ainda lembrar que, quanto à extração dos elementos, as amostras podem ser :

- Com reposição - quando um elemento sorteado puder ser sorteado novamente.
- Sem reposição - quando o elemento sorteado puder figurar uma única vez na amostra.

Pesquisa "Experimental" x Pesquisa "Correlacional"

Há dois tipos de pesquisas empíricas:

- **correlacional**: em que o pesquisador não influencia (ou tenta não influenciar) nenhuma variável, mas apenas as mede e procura por relações (correlações) entre elas, como idade e tamanho.
- **experimental**: em que o pesquisador manipula algumas variáveis e então mede os efeitos desta manipulação em outras variáveis; por exemplo, aumentar artificialmente a pressão sanguínea e, então, registra o nível de colesterol.

Note-se que na análise de dados em uma **pesquisa experimental** também se calcula a correlação entre as variáveis manipuladas e as que foram afetadas pela manipulação. Assim, os dados experimentais podem demonstrar conclusivamente relações causais (ou seja, relações de causa e efeito) entre as variáveis. Por exemplo, se o pesquisador descobrir que sempre que muda a variável 1 a variável 2 também muda, então poderá concluir que 1 influencia 2.

Paralelamente, em termos causais os dados de uma **pesquisa correlacional** podem ser apenas "interpretados", com base em outras teorias (não estatísticas) que o pesquisador conheça, mas não podem conclusivamente provar causalidade.

Amostra

Uma amostra estatística consiste em um subconjunto da população, ou seja, em um conjunto de indivíduos retirados de uma população a fim de que seu estudo estatístico possa fornecer informações importantes sobre a mesma população.

Deve-se notar que os elementos amostrais podem ser:

- Simples (indivíduos) ou
- Coletivos (famílias, irmandades, colônias).

Para se fazer as observações individuais toma-se a **Menor Unidade Amostral**:

Amostra	M.U.A.	Observação
---------	--------	------------

		individual
1. Peso de 100 ratos	cada rato	Peso de cada rato
2. Peso de 1 rato durante 1 mês	o rato	Cada pesagem do rato
3. Temperatura de formigueiros	cada formigueiro	Cada medição da temperatura

Parâmetro

É um valor singular que caracteriza a população. Exemplo: A média da altura de alunos de uma certa turma é 1,72 cm.

Estimativa

É um valor aproximado do parâmetro, sendo calculado com o uso da amostra.

Atributo

Se os dados apresentam um caráter **qualitativo** o seu levantamento e estudo é designado genericamente de estatística de atributo. **Exemplo:** Classificação dos alunos do CCB quanto ao sexo (classificação dicotômica do atributo)

Variável

Propriedade em relação à qual os indivíduos de uma amostra diferem de algum modo aferível (aferir = ajustar a um padrão). Note-se que as propriedades que não variam não são de interesse estatístico.

Há muitos modos de dividir os diferentes tipos de variáveis.

Níveis de mensuração das variáveis

As variáveis diferem em "quão bem" elas podem ser medidas, ou seja, em quanta informação seu nível de mensuração pode gerar.

Operacionalmente, muitas vezes pode-se estudar algo de diferentes maneiras.

Por exemplo, supondo que pretende-se estudar o hábito de fumar. Qual seria a escala? Haveria apenas 2 grupos: fumantes e não fumantes? Ou seria contado o número de cigarros fumados em 1 período? Como seria definido o fumante? Alguém que fuma 1 cigarro por dia será considerado fumante? E alguém que fuma 1 maço de cigarros por dia? São pertencentes à mesma categoria?

Assim, de acordo com sua escala de medição, as variáveis podem ser classificadas em 3 tipos:

- **Nominais:** São qualitativas. Os dados podem ser distribuídos em categorias mutuamente exclusivas. Seus valores só são registrados como nomes, só permitindo classificação qualitativa, não existindo ordem entre as categorias existentes. Assim, pode-se dizer que 2 indivíduos são diferentes em termos da variável analisada, mas não se pode dizer qual deles "tem mais" da qualidade representada pela variável.

Exemplos: sexo, cor dos olhos, causa de morte, grupo sanguíneo, raça, doente/sadio, fumante/não fumante.

- Análise: estudo de proporções e teste de Qui-quadrado.

- **Ordinais:** São qualitativas. Os dados podem ser distribuídos em categorias mutuamente exclusivas, mas que têm ordenação natural. São aquelas com possíveis resultados nominais, sem valores métricos, mas em que existe uma ordenação entre as categorias, com um resultado precedendo o outro. Portanto, permitem ordenar os itens medidos em termos de qual tem menos e qual tem mais da qualidade representada pela variável, mas não possibilitam que se diga "o quanto mais".

Exemplos: estágio da doença (inicial, intermediário, terminal); escolaridade (1o, 2o, 3o graus); mês de observação (janeiro, fevereiro, ..., dezembro); seqüência de eclosão de pupas (1^a, 2^a, 3^a, etc); peso, quando medido em 3 níveis (leve, médio, pesado); nível sócio-econômico de famílias residentes em uma localidade (miserável, pobre, classe média, rica); estágio da doença (inicial, intermediário, final).

- Análise: estudo de proporções, medianas, quartis, moda. Testes: Qui-quadrado, Crustal-Wallis, regressão logística e outros testes não paramétricos.

- **Intervalares:** São quantitativas. Os dados são expressos por números. Permitem não apenas ordenar os itens que estão sendo medidos, mas também possibilitam quantificar e comparar o tamanho das diferenças entre eles. Seus valores são medidos em uma escala métrica..

Exemplos: Temperatura: pode-se dizer que a temperatura de 20°C é maior do que 10°C e que um aumento de 10°C para 20°C é **duas vezes maior** do que um aumento de 15°C para 20°C; idade, em anos; Peso corporal em quilos.

- Análise: estudo de proporções, medianas, quartis, moda. Teste: Qui-quadrado, Kruskal-Wallis. Também se pode estudar as médias, desvios-padrão, efetuar análise de variância, a correlação e regressão linear e outros testes não paramétricos.

É evidente que as variáveis quantitativas incluem mais informação, portanto permitem que sejam aplicadas provas estatísticas mais potentes.

Dependência

Os termos variável dependente e independente são principalmente aplicados à pesquisa experimental. Portanto, há variáveis "independentes" dos padrões de reação inicial, das intenções e das características dos sujeitos da pesquisa (unidades experimentais). E existem variáveis "dependentes" da manipulação ou das condições experimentais, ou seja, elas dependem "do que os sujeitos farão" em resposta.

Esses termos também são aplicados em estudos em que não se manipulam variáveis, mas apenas se designam sujeitos a "grupos experimentais" baseados em propriedades pré-existentes dos próprios sujeitos.

Exemplo: se for comparada a % de hemoglobina em células de homens e mulheres, sexo pode ser considerado como variável independente e a % de hemoglobina como a variável dependente.

Assim, pode-se pensar que as variáveis independentes são aquelas que são manipuladas, enquanto que as dependentes são apenas medidas ou registradas.

- **Independentes:** Variáveis introduzidas propositadamente, para verificar sua relação com o comportamento de outras variáveis

- **Dependentes ou Resposta:** Variáveis cujo comportamento se quer verificar em função das variáveis independentes. Portanto, todo resultado obtido em um experimento é uma variável Dependente.

- **Espúrias:** Não são objeto do estudo, mas interferem no resultado. Devem ser cuidadosamente controladas.

Efeito

- **Moderadoras:** auxiliam na ocorrência de determinado efeito (do mesmo jeito que as independentes), mas são consideradas secundárias.

- **Intervenientes:** – são variáveis que ampliam, diminuem ou anulam o efeito das variáveis independentes sobre as dependentes. Não podem ser controladas. As variáveis intervenientes são consideradas causas da dependente.

- **Antecedentes:** são as causas do problema que originou a pesquisa.

Qualidade ou quantidade

• Variáveis qualitativas

São também chamadas de **categóricas** ou de **atributos**. São características que representam uma classificação dos indivíduos em categorias. Portanto, são definidas por categorias. Seus valores são expressos por atributos: sexo (macho ou fêmea), cor do olho (verde ou azul ou castanho), grupos de um sistema sanguíneo ...

Uma variável qualitativa pode ser registrada de uma forma quantitativa. Por exemplo se o sexo do indivíduo for registrado como 0 se fêmea ou 1 se macho, isto não significa que a variável sexo passou a ser quantitativa. Mas, combinando-se os atributos com frequências, esses podem ser tratados estatisticamente.

• Variáveis quantitativas

Os dados têm caráter nitidamente quantitativo, ou seja, são numéricos.

Mas, é importante lembrar que nem todas as variáveis representadas por números são quantitativas. Exemplos: o número do RG de um pessoa, o número do telefone de uma pessoa, o número da casa de uma pessoa... o número de seu cartão de identidade. Às vezes o sexo do indivíduo é registrado na planilha de dados como 1 se macho e 2 se fêmea, por exemplo. Evidentemente, isto não significa que a variável sexo passou a ser quantitativa.

Ressalte-se também que uma variável originalmente quantitativa pode ser tratada qualitativamente. Por exemplo:

- variável idade (em anos completos): é quantitativa (contínua); mas, se for informada apenas a faixa etária (0 a 10 anos, 10 a 20 anos, etc...) é qualitativa (ordinal).

- variável peso de lutadores de boxe: é uma variável quantitativa (contínua) se for utilizado o valor obtido na balança. Entretanto será qualitativa (ordinal) se o peso for usado para classificar o boxeador nas categorias peso-pena, peso-leve, peso-pesado.

As **variáveis quantitativas** podem ser classificadas também de dois modos:

- **Contínuas**: Podem assumir qualquer valor. Dependem da precisão do instrumento de medição. Exemplos: comprimentos, áreas, volumes, pesos, ângulos, temperaturas, períodos de tempo, taxas, %...

- **Descontínuas**: Podem assumir apenas valores inteiros, sem intermediários possíveis. Também são chamadas de **merísticas** ou **discretas**. Exemplos: Número de estruturas (ex: dentes, glândulas), Número de colônias de microorganismos em uma placa de Petri, Número de indivíduos por área.

Dado

Elemento ou quantidade conhecida que serve de base à resolução de um problema.

É importante notar que

- para a **Estatística** não interessam resultados **únicos**.
- se os dados não puderem ser **quantificados** não poderão ser submetidos à análise estatística.

Os **dados** podem ser

- **Primários**: são publicados pela própria pessoa ou organização que os haja recolhido.
- **Secundários**: são publicados por outra organização.

A **coleta** de dados pode ser:

- **Direta**: quando é obtida diretamente da fonte. Se subdivide em

- **Contínua** Exemplo: registros de nascimento, casamentos, óbitos

- **Periódica** Exemplo: censo demográfico

- **Ocasional** Exemplo: registro de casos de uma certa doença

- **Indireta**: É feita por deduções a partir dos elementos obtidos pela coleta direta, por analogia, por avaliação, indícios ou proporcionalização.

Limites implícitos

O último dígito de uma medida deve ser significativo, ou seja, deve estar entre limites implícitos, os quais devem ser verificáveis. Por exemplo:

Se a medida é	os limites implícitos são	
12	11,5	12,5
12,3	12,25	12,35
12,32	12,315	12,325

Arredondamento: uma convenção

Se o dígito a ser arredondado é seguido de um número menor que cinco, não é alterado. Se o dígito a ser arredondado é seguido de um número maior ou igual a cinco, é acrescido de uma unidade.

Exemplo: Supondo que se deseja arredondar para

- a terceira casa decimal:

3,141289: **3,141** 4,26781: **4,268** 2,99251: ?

- a segunda casa decimal:

3,141289: **3,14** 4,26781: **4,27** 2,99251: ?



Copie esse texto em formato [pdf](#) clicando em pdf com o botão direito do *mouse*.

Depois, clique em algo semelhante a "Salvar destino como"
Escolha um *drive* e uma pasta e clique em OK.

Biometria

Início

Topo

Este "site", destinado prioritariamente aos alunos de Fátima Conti, está disponível sob FDL ([Free Documentation Licence](#)), pretende auxiliar quem se interessa por Bioestatística, estando em permanente construção.

Sugestões e comentários são bem vindos.

Se desejar colaborar clique [aqui](#). Agradeço antecipadamente.



Deseja **enviar** essa página?

Se você usa um programa de correio eletrônico devidamente configurado para um [e-mail pop3](#), clique em "Enviar página" (abaixo) para abrir o programa.

Preencha o endereço do destinatário da mensagem.

E pode acrescentar o que quiser.

(Se não der certo, clique [aqui](#) para saber mais).

[Enviar página](#)



Se você usa [webmail](#) copie o endereço abaixo

<http://www.cultura.ufpa.br/dicas/biome/biodavar.htm>

Acesse a página do seu provedor. Abra uma nova mensagem.

Cole o endereço no campo de texto.

Preencha o endereço do destinatário.

E também pode acrescentar o que quiser.

Última alteração: 18 ago 2007