

O Processo Estatístico: tipos de variáveis

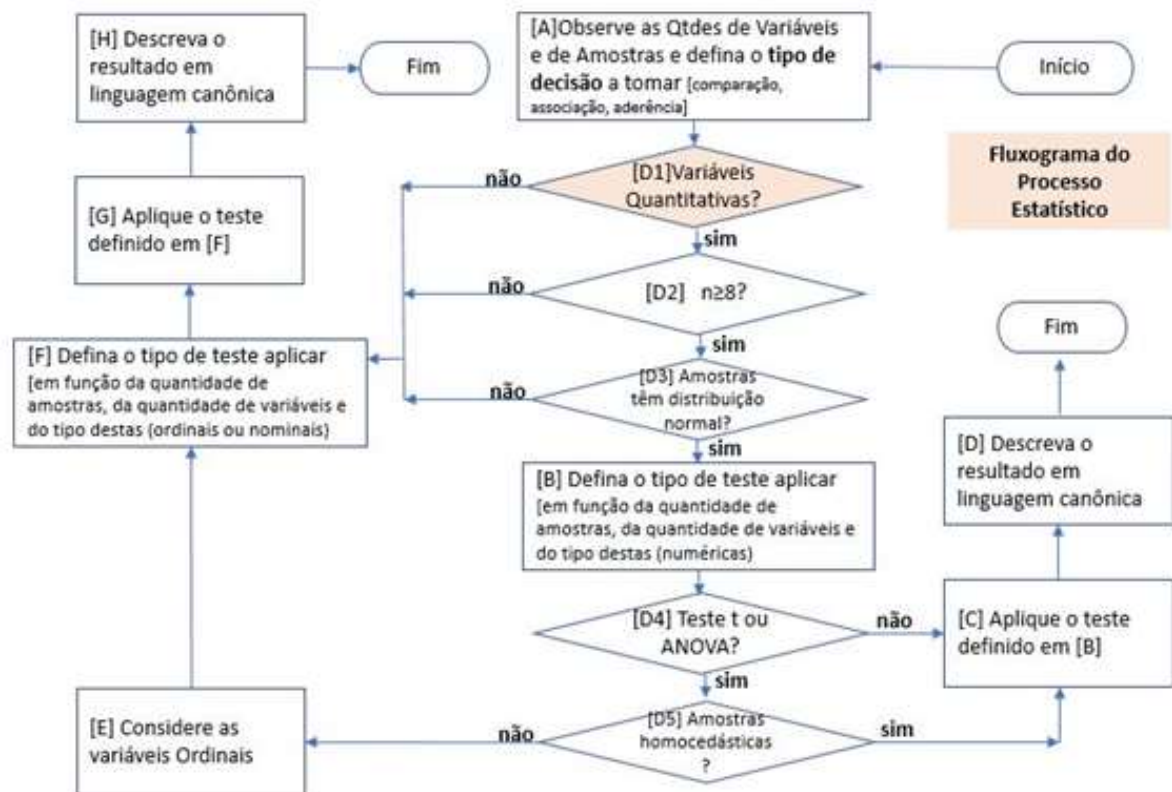
Manuel Meireles
UNIFACCAMP
meireles@faccamp.br

Este trabalho é continuação do publicado na edição anterior e tem por objetivo discorrer sobre o conceito de variável e definir seu tipo de forma correta, pois isso é vital para a boa condução do processo estatístico.

A análise estatística (de qualquer problema estatístico) requer algumas etapas:
1-reconhecer o tipo de decisão a fazer;
2-reconhecer os tipos de variáveis envolvidas;
3-definir o teste a aplicar;
4-aplicar o teste;
5-interpretar o resultado pelo p-value;
6-redigir a análise em linguagem canônica, isto é, em linguagem precisa dentro do estilo de linguagem da comunicação estatística.

Este trabalho ocupa-se da etapa 2 já que a etapa 1 foi objeto de estudo no número anterior.

Figura 1: O resultado da decisão <D1> leva a tipos de testes diferentes



1 Tipos de variáveis

Depois de sabido o tipo de decisão a tomar (se aderência, comparação ou associação) o Processo Estatístico solicita que se saiba com que tipos de variáveis o problema lida: é a decisão <D1> dentro do processo estatístico. Basicamente a decisão <D1> do Processo Estatístico requer que se responda à questão: “as variáveis envolvidas no problema são quantitativas?” O BioEstat considera três tipos de variáveis, ou três tipos de dados: quantitativos, ordenativos e categóricos. É preciso reconhecer estes tipos de mensuração das variáveis ou, simplificada e falando, os tipos de variáveis. Ver Figura 1.

Para entendermos os tipos de variáveis precisamos entender o conceito de variável, mas, entender este conceito é necessário ter em mente o conceito de objeto e é por aqui que se começa.

2 Conceito de objeto

Para fazer análise de dados é preciso saber não só quais as variáveis envolvidas, mas também a tipologia de tais variáveis. Antes de nos debruçarmos sobre as propriedades das variáveis devemos entender que as variáveis de uma pesquisa são decorrentes do objeto que ela estuda. Por **objeto deve-se entender algo tangível ou intangível passível de ser descrito** (Durkheim, 1919).

Um objeto, no sentido mais comum é tudo o que é manipulável ou manufaturável; é qualquer coisa, é qualquer peça ou artigo. O objeto é, existe, independente de nós e uma descrição do objeto, tal como ele é, é dita uma descrição objetiva. Uma descrição objetiva pode descrever um fato, um fenômeno.

O conceito ampliado de objeto abarca não apenas coisas materiais, mas também fatos, fenômenos, eventos. Durkheim (1919), nas suas *Regras do método sociológico* introduziu o conceito de **fato social** como sendo uma coisa. Para Durkheim, tratar certos fatos como coisas é ter para com eles uma certa atitude mental; é abordar o seu estudo partindo do princípio de que se desconhecem por completo e que as suas propriedades características, tal como as causas de que dependem, não podem ser descobertas pela introspeção, por mais atenta que seja. Um objeto é, desta forma, algo tangível ou intangível passível de ser observado e descrito.

Atributos do objeto

Para se descrever adequadamente um objeto se faz uso de atributos, obviamente atributos que sejam adequados à sua descrição. O termo atributo vem do latim *attributio* com o sentido de propriedade característica de alguém ou de alguma coisa. O atributo é um predicado objeto. Praticamente todos os atributos ou sub-atributos de objetos, dentro das organizações cumprem a função de descritores. Assim é comum a designação de um pelo outro e o conceito de descritor passa a envolver o de atributo.

Os descritores de um objeto descrevem o objeto dentro de um determinado contexto. Por exemplo, Maria pode ser descrita com os descritores apropriados como namorada de Jorge, ou descrita como estudante do curso de Medicina, ou como enfermeira de um Hospital. Desta forma diz-se que um objeto é descrito pelos seus **descritores funcionais** isto é, descritores de um objeto associados a um dado contexto.

Ter descritores funcionais não é suficiente: é necessário que a eles sejam atribuídos valores, medidas, pois só as medidas permitem comparações e o gerenciamento. Mensuração de métrica pode ser definida como atribuição de medida a um descritor funcional.

Mensuração é a atribuição de um valor a um descritor funcional de um objeto. Seja o objeto “aluno” que pode ser descrito por alguns descritores funcionais: idade; altura; sexo; série acadêmica e desempenho acadêmico. Ver exemplo no Quadro 1.

Quadro 1: Descritores Funcionais e respectivas métricas de Aluno

Objeto= ALUNO	
DESCRITOR FUNCIONAL	MÉTRICA
Idade	27
Altura	1.67
Sexo	0
Série acadêmica	3
Desempenho acadêmico	2

3 Conceito de variável

Pode-se entender variável como conjunto de um tipo de descritor funcional de uma classe de objetos associados às respectivas medidas. Variável é, portanto, o conjunto de valores atribuídos a um tipo de descritor de determinado objeto. A variável idade, por exemplo, inclui todos os valores observados referentes ao descritor Idade do objeto Aluno. Ver Quadro 2 como exemplo.

Quadro 2: Descritores do Objeto Aluno

Nome	Ana	João	Tiago
Idade	27	35	22
Altura	1.67	1.74	1.75
Sexo	0	1	1
Série acadêmica	3	4	1
Desempenho acadêmico	B	A	D

O Quadro 3 explicita o conceito de variável como conjunto de um tipo de descritor funcional (idade) de uma classe de objetos (aluno) associados às respectivas medidas (métricas).

Quadro 3: Variável e métricas correspondentes ao objeto

Amostra = Alunos		
Objeto	Descritor funcional (Variável)	Métrica
Ana	Idade	27
João	Idade	35
Tiago	Idade	22
Manuel	Idade.	29
Odair	Idade	26
Pedro	Idade	24

As variáveis são classificadas de acordo com o tipo de métrica utilizada para as dimensionar. É preciso que isto fique bem claro: por exemplo é possível expressar a idade como sendo uma variável de razão, ou intervalar, ordinal ou nominal. Vejamos como:

- de razão (numérica): Ana tem 27.42 (ou 27 anos completos)
- intervalar (numérica): Ana nasceu em 1994
- ordinal (ordenativa): Ana está na 3^a década
- nominal (categórica): Ana é adulta

As métricas são medidas quantificáveis e neste sentido podem ser entendidas como equivalentes a dados. Assis et al, (2019) chamam a atenção para este fato no verbete:

DADOS: Conjunto de qualquer tipo de informação detalhada e quantificada, resultado de medições ou experiências realizadas com objetivos específicos, usado como referência para determinações, estudos e trabalhos científicos. Toda a informação factível de ser resumida em um código, uma cifra, um esquema, um plano ou uma foto. Quer dizer, informação que não requer um texto ou um comentário para ser inteligível ou utilizável.

Quadro 4: Descritor funcional e correspondentes tipos de métricas

Objeto= ALUNO JOSÉ		
Descritor funcional	Métrica	Tipo de Métrica
Idade	27	Quantitativa (numérica)
Altura	1.67	Quantitativa (numérica)
Sexo	0	Nominal (categórica)
Série acadêmica	3	Ordinal (ordenativa)
Desempenho acadêmico	2	Ordinal (ordenativa)

O Quadro 4 está relacionado ao objeto “José” no contexto funcional de “Aluno” e para tal apresenta cinco descritores funcionais com as respectivas métricas, ou seja, indicadores cujos tipos estão indicados.

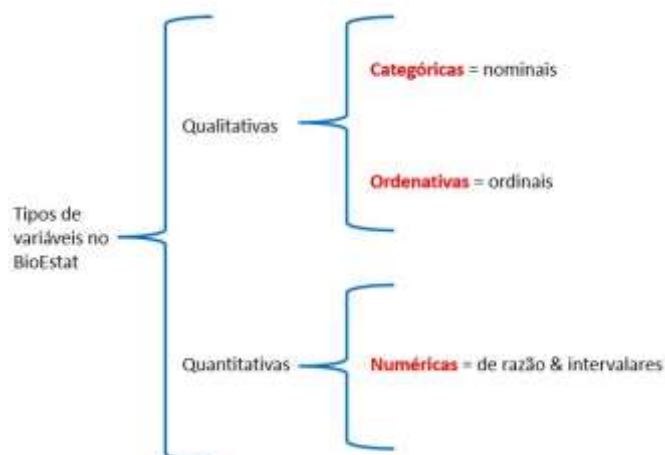
Cabe destacar que as métricas são tabuladas usualmente de forma numérica, pelo que idade (em anos), altura (em metros) são facilmente entendidos os valores. Já sexo (masculino ou feminino) e desempenho acadêmico (A, B, C e D) usam o que se denomina variáveis dummy: variáveis que usam artificialmente um dígito (ou número) para representar uma categoria. Por exemplo: Sexo Masculino= 1; sexo feminino =0; desempenho acadêmico: A=1, B=2, etc.

5–Níveis de mensuração

Há fundamentalmente dois grandes grupos de medidas: qualitativas e quantitativas. As qualitativas podem ser: nominais e ordinais e as quantitativas: intervalares e de razão.

A tipologia das variáveis no BioEstat é mais simplificada, pois considera as Categóricas (nominais) as Ordenativas (ordinais) e as Numéricas que são as variáveis quantitativas (as intervalares e de razão). Ver Figura 2.

Figura 2: Tipologia de variáveis adotada pelo BioEstat



Essencialmente a etapa 2 do Processo Estatístico é reconhecer os **tipos de variáveis** envolvidos isto é, observar se as variáveis envolvidas são categóricas, ordenativas ou numéricas.

Nominais ou Categóricas:

A mensuração *nominal* refere-se apenas a características nominativas descritoras do objeto. É o mais baixo nível de mensuração. São apenas atribuídos nomes aos descritores do objeto. As mensurações nominativas permitem apenas a relação de equivalência (=) que é reflexiva, simétrica e transitiva:

Reflexiva: $x=x$

Simétrica: se $x=y$, então $y=x$

Transitiva: se $x=y$ e $y=z$, então $x=z$

Exemplos de mensuração nominal ou categórica:

Diagnóstico psiquiátrico: Esquizofrênico; Paranóico; Maníaco-depressivo; Psiconeurótico

Sexo: Masculino; Feminino

Nacionalidade: Argentina; Brasileira; Chilena

Cor: Branca; Preta; Azul

Estado: Sim; Não

Ordinais ou Ordenativas (ou por postos)

Neste nível de mensuração é possível estabelecer certo tipo de relação lógica e natural (estabelecer uma ordem de preferência indicada pelo símbolo $>$).

As mensurações *ordinais* permitem a relação de equivalência (=) e a relação de comparação ($>$). Esta última é irreflexiva, assimétrica e transitiva, isto é:

Irreflexiva: não é verdade que para qualquer x se tenha $x>x$;

Assimétrica: se $x\geq y$, então $y\leq x$;

Transitiva: se $x=y$ e $y=z$, então $x=z$.

Exemplos de mensuração ordinal ou ordenativa:

Desempenho acadêmico: Conceito A; Conceito B; Conceito C

Preferência em escala: Concordo totalmente; Concordo; Indiferente; Discordo; Discordo totalmente

Status sócio- econômico: Classe Alta; Classe Média; Classe Baixa

Graduação militar: Sargento; Cabo; Soldado

Intervalar

É possível estabelecer certo tipo de relação (estabelecer uma ordem de preferência indicada pelo símbolo $>$) e, além disso se conhece as distâncias entre dois números quaisquer da escala. Neste tipo de mensuração não só se conhece a ordem como se estabelece que há intervalos iguais na escala.

Exemplos de mensuração intervalar:

Ano ou década de evento (o ano zero é arbitrário)

Temperatura em $^{\circ}\text{C}$ ou $^{\circ}\text{F}$ (a unidade e o ponto zero são arbitrários)

Uma característica deste tipo de mensuração é que não permite afirmar que o dobro de um valor represente efetivamente duas vezes mais tal valor. Por exemplo, não se pode afirmar que quem nasceu na década de 80 tenha alguma propriedade em dobro de quem nasceu na década de 40; um corpo à temperatura de 60°C não tem o dobro do calor de um corpo a 30°C .

Esta é a primeira escala verdadeiramente quantitativa e sobre variáveis intervalares é

possível aplicar todos os estatísticos paramétricos comuns. Com efeito, médias, desvios-padrão, etc são aplicáveis a dados em uma escala intervalar, bem como as provas paramétricas comuns (Testes t e F).

De razão ou Numéricas

As mensurações em escalas de razão, possuem todas as propriedades de uma escala intervalar e um verdadeiro ponto zero como origem. Todas as quatro relações seguintes são operacionalmente passíveis de ser obtidas:

Equivalência;

Comparação (>);

Razão entre dois intervalos;

Razão conhecida de dois valores quaisquer da escala.

A estas mensurações aplicam-se todos os estatísticos paramétricos comuns apontados para a escala intervalar além da média geométrica e o coeficiente de variação.

Davis (1976) chama atenção para o fato de que as propriedades da escala são cumulativas, isto é: as escalas mais complexas possuem todas as propriedades das menos sofisticadas e mais alguma coisa.

Davis (1976:24) oferece a tabela exibida na Figura 3.

Figura 3: Propriedades dos tipos de variáveis

propriedade	Escala complexidade →				
	Nominal	Ordinal	de intervalo	de razão	
1. Classificação	☀	☀	☀	☀	
2. Ordenação		☀	☀	☀	
3. Distância			☀	☀	
4. Zero não-arbitrado				☀	
		também conhecidas por categóricas		também conhecidas por escalares (scale)	

As decisões <D2> e <D3> são relativamente fáceis de tomar e serão abordadas no próximo número.

Referências

Assis, J. P.; Sousa, R. P., Dias, C.T.S. (2019). Glossário de estatística. Mossoró: EdUFERSA, 2019.

BioEstat <https://www.mamiraua.org.br/downloads/programas/>

DAVIS, James. Levantamento de dados em sociologia. Rio de Janeiro: Zahar, 1976.

Durkheim, É. (1919). Les règles de la méthode sociologique. Paris: Félix Alcan.