Aula 01 – Introdução à Estatística Descritiva – Tabelas

Luciana Rocha Pedro

GCC 1518 – Estatística e Probabilidade – CEFET Maracanã

USP – Projeto de Ensino

Aprender Fazendo Estatística

Estatística

A Estatística é uma ciência cujo campo de aplicação estende-se a muitas áreas do conhecimento humano.

Entretanto, um equívoco comum que deparamos nos dias atuais é que muitos pesquisadores consideram-se aptos a fazerem análises e inferências estatísticas sem um conhecimento mais aprofundado dos conceitos e teorias.

Tal prática, em geral, culmina em interpretações equivocadas e muitas vezes errôneas.

Estatística

Em sua essência, a Estatística é a ciência que apresenta processos próprios para **coletar, apresentar e interpretar** adequadamente **conjuntos de dados**, sejam eles numéricos ou não.

Podemos dizer que o seu objetivo é o de apresentar informações sobre dados em análise para que se tenha maior compreensão dos fatos que os mesmos representam.

Estatística Descritiva, Probabilística e Inferencial

A Estatística subdivide-se em três áreas: descritiva, probabilística e inferencial.

Estatística Descritiva

A estatística descritiva, como o próprio nome já diz, se preocupa em descrever os dados.

Estatística Probabilística

A estatística probabilística nos permite descrever os fenômenos aleatórios, ou seja, aqueles em que está presente a incerteza.

Estatística Inferencial

A estatística inferencial, fundamentada na teoria das probabilidades, se preocupa com a análise destes dados e sua interpretação.

Estatística Descritiva

A estatística descritiva, cujo objetivo básico é o de sintetizar uma série de valores de mesma natureza, organiza e descreve os dados de três maneiras: por meio de **tabelas**, **gráficos** e **medidas descritivas**.

Uma tabela é um quadro que resume um conjunto de observações, enquanto que gráficos são formas de apresentação dos dados, cujo objetivo é o de produzir uma impressão mais rápida e viva do fenômeno em estudo.

Medidas Descritivas

Para ressaltar as tendências características observadas nas tabelas, isoladamente, ou em comparação com outras, é necessário expressar tais tendências através de números ou estatísticas.

Estes números ou estatísticas, também chamados de medidas descritivas, são divididos em duas categorias: **medidas de posição** e **medidas de dispersão**.

População

A estatística trabalha com dados, os quais podem ser obtidos por meio de uma **população** ou de uma **amostra**.

Definição

População: conjunto de elementos que tem pelo menos uma característica em comum.

Esta característica deve delimitar corretamente quais são os elementos da população, que podem ser animados ou inanimados.

Amostra

Definição

Amostra: subconjunto de elementos de uma população.

Este subconjunto deve ter dimensão menor que o da população e seus elementos devem ser representativos da população.

A seleção dos elementos que irão compor a amostra pode ser feita de várias maneiras e irá depender do conhecimento que se tem da população e da quantidade de recursos disponíveis.

A estatística inferencial é a área que trata e apresenta a metodologia de amostragem.

Conjuntos Finitos e Infinitos

Em se tratando de conjuntos (e subconjuntos), estes podem ser:

Definição

Conjuntos Finitos: possuem um número limitado de elementos.

Definição

Conjuntos Infinitos: possuem um número ilimitado de elementos.

Neste contexto, elemento significa cada uma das unidades observadas no estudo.

Variável

Após a determinação dos elementos, perguntamos: o que fazer com eles? Podemos medi-los, observá-los e contá-los, gerando um conjunto de respostas que receberá a denominação de variável.

Definição

Variável: é a característica que vai ser observada, medida ou contada nos elementos da população ou da amostra e que pode variar, ou seja, assumir um valor diferente de elemento para elemento.

Tipos de Variáveis

Não basta identificar a variável a ser trabalhada, é necessário fazermos a distinção entre os tipos de variáveis:

Definição

Variável qualitativa: é uma variável que assume como possíveis valores atributos ou qualidades. Também são denominadas variáveis categóricas.

Definição

Variável quantitativa: é uma variável que assume como possíveis valores, números.

Variável Qualitativa

Uma variável qualitativa pode ter duas classificações:

Definição

Variável qualitativa nominal: é uma variável que assume como possíveis valores atributos ou qualidades e estes não apresentam uma ordem natural de ocorrência.

Definição

Variável qualitativa ordinal: é uma variável que assume como possíveis valores atributos ou qualidades e estes apresentam uma ordem natural de ocorrência.

Variável Quantitativa

Uma variável quantitativa também pode ter duas classificações:

Definição

Variável quantitativa discreta: é uma variável que assume como possíveis valores números, em geral inteiros, formando um conjunto finito ou enumerável.

Definição

Variável quantitativa contínua: é uma variável que assume como possíveis valores números, em intervalos da reta real e, em geral, resultantes de mensurações.

Exemplos

Classifique os exemplos a seguir:

1. Número de reprovações, por disciplina, dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: $0, 1, 2, \ldots$

2. Estado civil dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: solteiro, casado, separado.

Exemplos

Classifique os exemplos a seguir:

1. Número de reprovações, por disciplina, dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: $0, 1, 2, \ldots$
   R: Variável quantitativa discreta

2. Estado civil dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: solteiro, casado, separado.

Exemplos

Classifique os exemplos a seguir:

1. Número de reprovações, por disciplina, dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: $0, 1, 2, \ldots$
   R: Variável quantitativa discreta

2. Estado civil dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: solteiro, casado, separado.
   R: Variável qualitativa ordinal

Exemplos

Classifique os exemplos a seguir:

3. Peso (quilogramas) dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: $58, 59, 63, \ldots$

4. Meios de informação utilizados pelos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: televisão, revista, internet, jornal.

Exemplos

Classifique os exemplos a seguir:

3. Peso (quilogramas) dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: $58, 59, 63, \ldots$
   R: Variável quantitativa contínua

4. Meios de informação utilizados pelos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: televisão, revista, internet, jornal.

Exemplos

Classifique os exemplos a seguir:

3. Peso (quilogramas) dos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: $58, 59, 63, \ldots$
   R: Variável quantitativa contínua

4. Meios de informação utilizados pelos alunos da disciplina Inferência Estatística do curso de Estatística da UEM: televisão, revista, internet, jornal.
   R: Variável qualitativa nominal

Dados

É muito comum nos dias de hoje, devido ao uso de computadores, realizarmos pesquisas em que a coleta de dados resulta em grandes quantidades de dados para análise e torna-se quase impossível entendê-los, se estes dados não estiverem resumidos.

Em outras palavras, os dados, na forma em que foram coletados, não permitem, de maneira fácil e rápida, que se extraia informações. Torna-se difícil detectar a existência de algum padrão.

É necessário trabalhar os dados para transformá-los em informações, para compará-los com outros resultados, ou ainda para julgar sua adequação a alguma teoria.

Exemplo – Tabela 01

Um questionário foi aplicado aos alunos de Estatística da Universidade Estadual de Maringá (UEM). As variáveis são:

- **Sexo**: sexo, com categorias (1) masculino e (2) feminino
- **Id**: idade, em anos
- **Altura**: altura, em metros e centímetros
- **Peso**: peso, em quilos
- **Est.Civil**: estado civil, com categorias (1) solteiro, (2) casado e (3) separado
- **N.Ir.**: número de irmãos
- **Transp.**: meio de transporte mais utilizado, com categorias (1) coletivo e (2) próprio
- **Procedência**: município de procedência, com categorias (1) Maringá, (2) outro município do Paraná e (3) outro Estado
- **Trabalho**: relação do trabalho com o curso, com categorias (1) não trabalho, (2) completamente relacionado, (3) parcialmente relacionado e (4) não relacionado
- **Inform**: meio de informação mais utilizado, com categorias (1) TV, (2) jornal, (3) rádio, (4) revista e (5) internet
- **Disc.**: número de disciplinas reprovadas no primeiro ano da UEM

Tabela 01

Para trabalharmos com estes dados são necessários, em primeiro lugar, tabulá-los e apresentá-los na forma em que foram coletados (dados brutos), como na Tabela 01.

Em geral, a primeira coluna da tabela deve conter a identificação do respondente.

Tabela 01

Tabela 01 – Informações sobre os alunos da disciplina Inferência Estatística do curso de Estatística da UEM – 21/03/2005.

Num	Sexo	Id	Altura	Peso	Est.Civil	N.Ir.	Transp.	Procedência	Trabalho	Inform	Disc.
1	F	20	1,60	58	Solteiro	1	Próprio	Maringá	Não rel.	TV	2
2	F	26	1,65	59	Solteiro	2	Coletivo	Fora do PR	Não trab.	Revista	0
3	F	18	1,64	55	Solteiro	2	Próprio	Maringá	Não trab.	TV	0
4	F	25	1,73	60	Solteiro	2	Coletivo	Outro no PR	Não rel.	TV	2
5	M	35	1,76	83	Casado	6	Coletivo	Outro no PR	Não rel.	TV	2
6	F	20	1,62	58	Solteiro	2	Coletivo	Outro no PR	Não rel.	Rádio	5
7	F	29	1,72	70	Solteiro	3	Coletivo	Maringá	Não trab.	TV	0
8	M	23	1,71	62	Separado	2	Próprio	Outro no PR	Não rel.	Internet	2
9	F	20	1,63	63	Solteiro	2	Próprio	Maringá	Não trab.	TV	1
10	M	20	1,79	75	Solteiro	2	Próprio	Fora do PR	Não trab.	Internet	2
11	M	20	1,82	66	Solteiro	1	Próprio	Fora do PR	Não trab.	TV	2
12	F	30	1,68	46	Solteiro	3	Próprio	Outro no PR	Parc.rel.	TV	4
13	F	18	1,69	64	Solteiro	1	Próprio	Maringá	Parc.rel.	TV	0
14	M	37	1,82	80	Casado	2	Próprio	Maringá	Não rel.	TV	3
15	M	25	1,83	62	Solteiro	1	Próprio	Outro no PR	Não rel.	TV	2
16	F	20	1,63	68	Solteiro	2	Coletivo	Maringá	Não trab.	TV	2
17	M	21	1,71	80	Solteiro	2	Coletivo	Maringá	Não rel.	Internet	0
18	M	25	1,80	82	Casado	1	Próprio	Outro no PR	Não rel.	Internet	3
19	F	24	1,62	55	Solteiro	2	Próprio	Maringá	Não trab.	Jornal	2
20	M	19	1,74	58	Solteiro	2	Próprio	Maringá	Com.rel.	TV	3
21	F	21	1,55	65	Solteiro	1	Próprio	Maringá	Não trab.	TV	1
22	M	22	1,73	62	Solteiro	0	Próprio	Maringá	Não trab.	Jornal	4

Observações

Podemos observar que a tabela de dados brutos contém muita informação, porém pode não ser muito rápido e prático obter estas informações.

Por exemplo, não é imediato afirmar que existem mais homens que mulheres.

Podemos, então, construir outra tabela para cada uma das variáveis, que resumirá as informações ali contidas.

Observamos também que, ao usarmos programas computacionais, associamos valores numéricos às variáveis qualitativas e nem por isso a variável deixa de ser qualitativa. Cabe ao bom senso lembrar da natureza da variável.

Tipos de Tabelas

Embora um certo volume de informação seja perdido quando os dados são resumidos, um grande volume pode também ser ganho.

Todas as variáveis podem ser resumidas através de uma tabela, mas a construção é diferenciada dependendo do tipo de variável.

Denominamos **tabela simples** a tabela que resume os dados de uma única variável qualitativa e **distribuição de freqüências** ao resumo de uma única variável quantitativa.

Tabela Simples

Uma tabela simples contém as diferentes categorias observadas de uma **variável qualitativa** e suas respectivas contagens, denominadas **freqüências absolutas**.

A contagem refere-se ao número de ocorrências de cada categoria e é realizada utilizando-se, por exemplo, uma tabela ou o banco de dados.

Classificação de uma Tabela Simples

Quanto à classificação, uma tabela simples pode ser:

- **temporal**, quando as observações são feitas levando-se em consideração o tempo;
- **geográfica**, quando os dados referem-se ao local de ocorrência;
- **específica** (ou categórica), quando tempo e local são fixos; e
- **comparativa**, quando a tabela resume informações de duas ou mais variáveis. A tabela comparativa é também denominada **tabela cruzada** ou de dupla ou mais entradas.

Exemplo

Exemplo de uma tabela histórica.

Tabela 02 – Número de alunos matriculados na disciplina Probabilidade I do curso de Estatística da Universidade Estadual de Maringá.

Ano	Nº de Alunos
2000	40
2001	59
2002	63
2003	69
2004	71

Fonte: DES/UEM.
Nota: Os números de 2003 e 2004 correspondem a duas turmas.

Exemplo

Exemplo de uma tabela geográfica, específica e comparativa, construída a partir da Tabela 01.

Tabela 03 – Município de procedência dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Município de Procedência	Nº de Alunos
Maringá	12
Outro no Paraná	7
Fora do Paraná	3
Total	22

Fonte: Tabela 01.

Freqüências Relativas e/ou Relativas em Percentual

É comum e útil na interpretação de tabelas a inclusão de uma coluna contendo as freqüências relativas e/ou relativas em percentual.

A **freqüência relativa** é obtida dividindo-se a freqüência absoluta de cada categoria da variável pelo número total de observações (número de elementos da amostra ou da população).

Multiplicando este resultado por 100, obtemos a **freqüência relativa em percentual**.

Freqüências Relativas e/ou Relativas em Percentual

As freqüências relativas em percentual são úteis ao se comparar tabelas ou pesquisas diferentes.

Por exemplo, quando amostras (ou populações) têm números de elementos diferentes, a comparação através das freqüências absolutas pode resultar em afirmações errôneas, enquanto que pelas freqüências relativas em percentual os percentuais totais são os mesmos.

Tabela com Freqüências Relativas em Percentual

Tabela 04 – Município de procedência dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Município de Procedência	Nº de Alunos	Percentual
Maringá	12	55
Outro no Paraná	7	32
Fora do Paraná	3	13
Total	22	100

Fonte: Tabela 01.

Exercícios

Construa uma tabela para o meio de transporte mais utilizado pelos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, de acordo com a Tabela 01.

Exercícios

Construa uma tabela para o meio de transporte mais utilizado pelos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, de acordo com a Tabela 01.

Tabela 05 – Meio de transporte mais utilizado pelos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Meio de transporte	Nº de Alunos
Coletivo	7
Próprio	15
Total	22

Fonte: Tabela 01.

Exercícios

Construa uma tabela para o meio de transporte mais utilizado segundo o sexo dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, de acordo com a Tabela 01.

Exercícios

Construa uma tabela para o meio de transporte mais utilizado segundo o sexo dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, de acordo com a Tabela 01.

Tabela 06 – Meio de transporte mais utilizado segundo o sexo dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Meio de transporte	Sexo		Total
	Masculino	Feminino	
Coletivo	2	5	7
Próprio	8	7	15
Total	10	12	22

Fonte: Tabela 01.

Exercício 01

Construa tabelas simples, incluindo os percentuais, para as variáveis estado civil, relação do trabalho com o curso de graduação e meio de transporte mais utilizado referentes à Tabela 01.

Construa, também, uma tabela cruzada para as variáveis estado civil e meio de informação.

Rol

Como já foi mencionado, dependendo do volume de dados, torna-se difícil ou impraticável tirar conclusões a respeito do comportamento das variáveis e, em particular, de variáveis quantitativas.

Podemos, no entanto, colocar os dados brutos de cada uma das **variáveis quantitativas** em uma ordem crescente ou decrescente, denominado **rol**.

A visualização de algum padrão ou comportamento continua sendo difícil, mas torna-se rápido identificar menores e maiores valores ou concentrações de valores.

Amplitude Total

Estes números (menor e maior valor observado) servem de ponto de partida para a construção de tabelas para estas variáveis.

É a diferença entre o menor e maior valor observado da variável X, denominada **amplitude total** ($AT = x_{\max} - x_{\min}$), que definirá a construção de uma **distribuição de freqüência pontual ou em classes**.

Para as variáveis qualitativas, podemos também construir um rol em ordem temporal ou alfabética, por exemplo.

O ideal é que uma distribuição de freqüência resuma os dados em um número de linhas que varie de 5 a 10.

Distribuição de Freqüência Pontual – Sem Perda de Informação

A construção de uma **distribuição de freqüência pontual** é equivalente à construção de uma tabela simples, em que listamos os diferentes valores observados da variável, com suas **freqüências absolutas**, denotadas por F_i, em que o índice i corresponde ao número de linhas da tabela.

Tabela 07 – Número de irmãos dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Número de irmãos	Contagem	Freqüência (F_i)
0	|	1
1	|||||	6
2	||||||||||	12
3	||	2
6	|	1
Total		22

Fonte: Tabela 01.

Observações

Observamos que esta variável foi resumida em 5 linhas. Assim, $i = 1, \ldots, 5$, e, portanto, temos cinco valores para as freqüências absolutas.

A freqüência absoluta da segunda linha, $F_2 = 6$, por exemplo, indica que seis alunos têm um irmão, enquanto apenas um afirmou ter seis irmãos.

A soma de todas as freqüências absolutas deve ser igual ao número total de observações da variável, neste caso, 22.

A segunda coluna desta tabela é uma coluna opcional em distribuições de freqüências.

Freqüência Relativa

Considerando as colunas complementares em uma distribuição de freqüências e i a ordem da linha na tabela, temos:

- a freqüência relativa, denotada por f_i, é definida como:

$$f_i = \frac{F_i}{n},$$

em que n é o tamanho da amostra, devendo ser substituída por N se os dados forem populacionais.

A soma das freqüências relativas de todas as categorias é igual a 1.

Freqüência Relativa em Percentual

- a freqüência relativa em percentual, denotada por $f_i\%$, é definida como:

$$f_i\% = \frac{F_i}{n} \cdot 100,$$

representa o percentual de observações que pertencem àquela categoria.

A soma das freqüências deve, agora, ser igual a 100%.

Freqüências Acumuladas

- a freqüência absoluta acumulada, denotada por F_{a_i}, é obtida somando-se a freqüência absoluta do valor considerado às freqüências absolutas anteriores a este mesmo valor.

- a freqüência relativa acumulada, denotada por $f_{a_i}\%$, é definida como:

$$f_{a_i}\% = \frac{F_{a_i}}{n} \cdot 100.$$

Distribuição de Freqüências Completa

Uma tabela contendo todas estas freqüências é dita uma **distribuição de freqüências completa**. Desta forma, a Tabela 8 pode ser apresentada como:

Tabela 08 – Número de irmãos dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Número de irmãos (x_i)	F_i	f_i %	F_{a_i}	f_{a_i} %
0	1	4,55	1	4,55
1	6	27,26	7	31,81
2	12	54,55	19	86,36
3	2	9,09	21	95,45
6	1	4,55	22	100,00
Total	22	100,00		

Fonte: Tabela 01.

Distribuição de Freqüência em Classes – Com Perda de Informação

A **distribuição de freqüências em classes** é apropriada para apresentar dados quantitativos contínuos ou discretos com um número elevado de possíveis valores.

É necessário dividir os dados em intervalos ou faixas de valores, que são denominadas **classes**.

Uma **classe** é uma linha da distribuição de freqüências. O menor valor da classe é denominado **limite inferior** (l_i) e o maior valor da classe é denominado **limite superior** (L_i).

Intervalos

O intervalo (ou classe) pode ser representado das seguintes maneiras:

1. $l_i|$_____ L_i, em que o limite inferior da classe é incluído na contagem da freqüência absoluta, mas o superior não;
2. l_i _____$|L_i$, em que o limite superior da classe é incluído na contagem, mas o inferior não;
3. $l_i|$_____$|L_i$, em que tanto o limite inferior quanto o superior são incluídos na contagem;
4. l_i _____ L_i, em que os limites não fazem parte da contagem.

Podemos escolher qualquer uma destas opções. O importante é tornar claro no texto ou na tabela qual está sendo usada.

Número de Classes

Geralmente, os seguintes critérios são utilizados para a determinação do **número de classes**, denotado por k:

- Raiz quadrada: $k = \sqrt{n}$,
- log (Sturges): $k = 1 + 3,3 \log n$,
- ln (Milone): $k = -1 + 2 \ln n$,
- $k = 1 + 10^d AT$,

em que n é o número de elementos da amostra, AT é a amplitude total dos dados e d é o número de decimais de seus elementos.

Devemos lembrar que, sendo k o número de classes, o resultado obtido por cada um dos critérios deve ser o número inteiro mais próximo ao obtido.

Amplitude de Cada Classe

Determinado o número de classes da distribuição de freqüências, o próximo passo é determinar **a amplitude de cada classe**, h, definida por:

$$h = \frac{AT}{k}.$$

Portanto, todas as classes terão a mesma amplitude, o que permitirá a construção de gráficos e o cálculo de medidas descritivas.

Pontos Médios

No caso de uma distribuição de freqüência contínua, ou em classes, uma outra coluna pode ser acrescentada à tabela. É a coluna dos **pontos médios**, denotada por x_i e definida como a média dos limites da classe:

$$x_i = \frac{l_i + L_i}{2}, \qquad i = 1, \ldots k.$$

Estes valores são utilizados na construção de gráficos e na obtenção de medidas descritivas com o auxílio de calculadoras.

Exemplo

Considere a variável idade dos alunos da Tabela 01. A Tabela 09 apresenta a distribuição de freqüência pontual.

Tabela 09 – Idade dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Idade	F_i
18	2
19	1
20	6
21	2
22	1
23	1
24	1
25	3
26	1
29	1
30	1
35	1
37	1
Total	22

Fonte: Tabela 01.

Distribuição de Freqüência em Classes

Podemos observar que a tabela possui 13 linhas e que muitas delas, seguidas, apresentam freqüência igual a 1, o que mostra que o resumo da idade não apresenta uma distribuição satisfatória dos dados.

Ao passar dos dados brutos para uma distribuição de freqüência em classes, algumas informações são perdidas, pois não temos mais as observações individuais.

Por outro lado, essa perda é pequena quando comparada ao ganho de concisão e de facilidade de interpretação da distribuição de freqüência.

Distribuição de Freqüência em Classes

Assim, para a idade, temos:

- $AT = 37 - 18 = 19$ anos;
- $k = \sqrt{22} = 4,69 \cong 5$ classes;
- $h = \dfrac{19}{5} = 3,8 \cong 4$ anos.

Distribuição de Freqüência em Classes

A distribuição de freqüência em classes é dada na Tabela 10.

Tabela 10 – Idade dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Idade	x_i	F_i	f_i %	F_{a_i}	f_{a_i} %
18 |---22	20	11	50,00	11	50,00
22 |---26	24	6	27,27	17	77,27
26 |---30	28	2	9,09	19	86,36
30 |---34	32	1	4,55	20	90,91
34 |---38	36	2	9,09	22	100,00
Total	-	22	100,00	-	-

Fonte: Tabela 01.

Observações

Notamos que cada um dos valores observados deve pertencer a uma e somente uma classe.

É usual que o limite inferior da primeira classe seja igual ao menor valor observado e que o maior valor pertença à última classe.

Quando o limite superior da última classe coincidir com o maior valor observado, é mais apropriado fechar este intervalo, contando o elemento nesta classe, do que abrir uma nova classe contendo apenas uma freqüência absoluta.

Por outro lado, se o maior valor observado for inferior ao limite superior da classe, não há problemas, pois fixamos todas as classes com a mesma amplitude.

Amplitudes de Classes Desiguais

Nada impede que construamos uma tabela com **amplitudes de classes desiguais**. Isto dependerá do objetivo do pesquisador. O que se recomenda é o cuidado na interpretação da tabela.

O primeiro passo é calcular as amplitudes das classes (Δ_i) e apresentá-las em uma coluna.

Em seguida, calculamos as densidades de freqüências de cada classe, dividindo F_i por Δ_i, para conhecermos a concentração por unidade da variável.

Podemos, também, calcular as densidades das proporções para conhecermos o percentual de concentração em cada classe (f_i/Δ_i).

Exemplo

Considere os dados do exemplo anterior. A distribuição de freqüências com intervalos de classes desiguais é apresentada na Tabela 11.

Tabela 11 – Idade dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá, 21/03/2005.

Idade	Frequência F_i	Amplitude Δ_i	Densidade F_i / Δ_i	Proporção f_i	Densidade f_i / Δ_i
18 |---20	3	2	1,50	0,14	0,07
20 |---22	9	2	4,50	0,40	0,20
22 |---24	2	2	1,00	0,09	0,05
24 |---28	5	4	1,25	0,23	0,06
28 |---38	3	10	0,30	0,14	0,01
Total	22	-	-	1,00	-

Fonte: Tabela 01.

Exemplo

Uma outra forma de construir uma distribuição de freqüências com amplitudes de classes desiguais é apresentada na Tabela 12, em que a última classe não apresenta limite superior especificado. Isto poderia, também, ocorrer na primeira classe, mas agora com o limite inferior não especificado.

Com este tipo de distribuição, dificuldades podem ocorrer na construção de gráficos e no cálculo da média, por exemplo.

Tabela 12 – Idade dos alunos da disciplina Inferência Estatística do curso de Estatística da Universidade Estadual de Maringá.

Idade	F_i	f_i %	Fa_i	fa_i %
18 |---20	3	14	3	14
20 |---22	8	36	11	50
22 |---24	2	9	13	59
24 |---26	4	18	17	77
Acima de 26	5	23	22	100
Total	22	100	-	-

Fonte: Tabela 01.

Exercício 02

Construa uma distribuição de freqüência completa para as variáveis da Tabela 01:

1. número de disciplinas reprovadas no primeiro ano do curso;
2. peso.