

ESTATÍSTICA

Professora: Patrícia Ferreira Paranaíba

Estatística Descritiva

A estatística descritiva é parte da estatística que lida com a organização, resumo e apresentação de dados. Esta é feita por meio de:

- Tabelas;
- Gráficos;
- Medidas Descritivas (média, variância, entre outras).

VARIÁVEIS

QUALITATIVAS (atributos)

Exemplos:

- Sexo;
- Religião;
- Naturalidade;
- Cor dos olhos;
- Altura de uma planta (baixa, média, alta);
- Cor de flor;
- Sabor:

NOMINAL
Ex: região;

ORDINAL
Ex: classe social;

QUANTITATIVAS (numéricas)

DISCRETAS

Exemplos:

- Quantidades de estudantes em uma disciplina;
- Quantidades de cômodos em uma residência;
- Número de filhos;

CONTÍNUAS

Exemplos:

- Tempo de voo entre cidades;
- Duração da bateria do celular;
- Peso corporal;

Variáveis Qualitativas

- Frequentemente o primeiro passo da descrição de dados é criar uma tabela de frequências. Antes de montar a tabela de distribuição de frequências temos algumas definições:
- Frequência - medida que quantifica a ocorrência dos valores de uma variável a um dado conjunto de dados. As frequências podem ser:
 - Absoluta (f_a) - contagem das observações de uma variável;
 - Relativa (fr) - divisão da frequência absoluta pelo total de observações

$$fr = \frac{f_a}{n}$$

- Percentual (fp) - é a frequência relativa multiplicada por 100

$$fp = 100 \times fr$$

Exemplo 1: Para adequar os produtos às preferências dos clientes, foi realizada uma pesquisa sobre os provedores e a qualidade dos serviços prestados utilizando uma amostra de 20 clientes, obtendo as seguintes variáveis:

Tabela: Variáveis observadas de 20 clientes de um provedor.

Amostra	Sexo	Qualidade	Amostra	Sexo	Qualidade
1	feminino	Boa	11	feminino	Ruim
2	feminino	Boa	12	feminino	Ruim
3	feminino	Boa	13	masculino	Boa
4	feminino	Boa	14	masculino	Boa
5	feminino	Boa	15	masculino	Ótimo
6	feminino	Ótimo	16	masculino	Regular
7	feminino	Ótimo	17	masculino	Regular
8	feminino	Regular	18	masculino	Ruim
9	feminino	Regular	19	masculino	Ruim
10	feminino	Ruim	20	masculino	Ruim

Para a variável sexo, podemos utilizar as frequências apresentadas na tabela 2:

Tabela: Distribuição de frequência do sexo de 20 clientes de um provedor.

Sexo	Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)
feminino	12	0,60	60%
masculino	8	0,40	40%
Total	20	1,00	100%

Para a variável qualidade no atendimento, além das frequências utilizadas para a variável sexo, podemos utilizar mais duas frequências:

- Frequência Acumulada (FA)- obtida pelo soma das frequências absolutas;
- Frequência Percentual Acumulada (FP) - obtida pela soma das frequências percentuais.

Tabela: Distribuição de frequência qualidade no atendimento de um provedor de acordo com 20 clientes

Qualidade no Atendimento	Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)	Frequência Acumulada (FA)	Frequência Percentual Acumulada (FP)
Ótima	3	0,15	15%	3	15%
Boa	7	0,35	35%	10	50%
Regular	4	0,20	20%	14	70%
Ruim	6	0,30	30%	20	100%
Total	20	1,00	100%	-	-

- Dados qualitativos são usualmente bem ilustrados num simples gráfico de barras onde a altura da barra é igual a frequência.

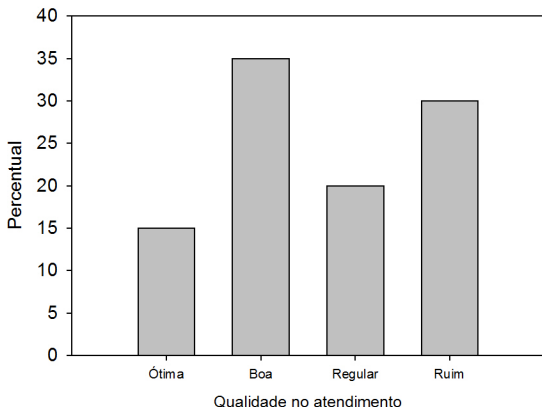


Figura: Qualidade no atendimento de um provedor de acordo com 20 clientes.

- Gráfico de Setor ou de Pizza.



Variáveis Quantitativas discretas

- Da mesma forma que as variáveis qualitativas, podemos resumir dados quantitativos por meio de tabelas de frequências.
- A tabela de distribuição de frequências de uma variável discreta é, em geral bastante semelhante à das variáveis qualitativas.

Exemplo 1: Os dados a seguirem são referentes a um levantamento onde observou-se o número de peças defeituosas em 25 máquinas de uma empresa.

Tabela: Número de peças defeituosas em 25 máquinas de uma empresa.

3	5	7	1	3
6	5	5	5	3
8	5	2	6	2
4	4	4	3	5
6	2	2	4	5

Tabela: Distribuição de frequências do número de peças defeituosas de 25 máquinas de uma empresa.

Número de Minerais	Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)	Frequência Acumulada (FA)	Frequência Percentual Acumulada (FP)
1	1	0,04	4%	1	4%
2	4	0,16	16%	5	20%
3	4	0,16	16%	9	36%
4	4	0,16	16%	13	52%
5	7	0,28	28%	20	80%
6	3	0,12	12%	23	92%
7	1	0,04	4%	24	96%
8	1	0,04	4%	25	100%
Total	25	1	100%		

Representação gráfica: gráfico de barras

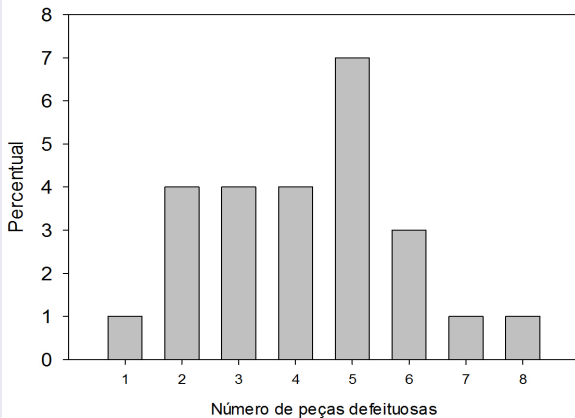


Figura: Número número de peças defeituosas de 25 máquinas de uma empresa.

Variáveis Quantitativas contínuas

- O procedimento para construir tabelas de distribuição de frequências para variáveis quantitativas contínuas envolve os seguintes passos :
 - Decidir sobre o número de classes k . Para que a decisão não seja totalmente arbitrária pode-se usar a raiz quadrada do total de valores como o número de classes, ou seja, $k \cong \sqrt{n}$.
 - Determinar a amplitude dos dados: $A = \text{Max} - \text{Min}$.
 - Determinar a amplitude de classe c :

$$c = \frac{A}{k - 1}$$

- Determinar o limite inferior da primeira classe LI_1 :

$$LI_1 = \text{Min} - \frac{c}{2}$$

- Determinar o limite superior da primeira classe LS_1 :

$$LS_1 = LI_1 + c$$

sendo que o limite inferior da segunda classe LI_2 é igual ao LS_1 , e assim

$$LS_2 = LI_2 + c$$

e assim, sucessivamente todas as classes vão sendo construídas.

- Após a construção das classes, são contados quantos dados estão contidos em cada classe e se obtêm as frequências.

Exemplo 2

Tabela: Dados ordenados, relativos ao tempo em segundos para carga de um aplicativo num sistema compartilhado (30 observações).

6,94	7,27	7,46	7,97	8,03	8,37
8,56	8,66	8,88	8,95	9,30	9,33
9,55	9,76	9,80	9,82	9,98	9,99
10,14	10,19	10,42	10,44	10,66	10,88
10,88	11,16	11,80	11,88	12,25	12,34

$$k = \sqrt{30} = 5,47 \approx 5$$

$$A = \text{Max} - \text{Min} = 12,34 - 6,94 = 5,40$$

$$c = \frac{A}{k-1} = \frac{5,40}{4} = 1,35$$

$$LI_1 = \text{Min} - \frac{c}{2} = 6,94 - \frac{1,35}{2} = 6,94 - 0,67 = 6,27$$

Tabela: Distribuição de frequências, relativa ao ao tempo em segundos para carga de um aplicativo num sistema compartilhado.

Classes			Frequência Absoluta (fa)	Frequência Relativa (fr)	Frequência Percentual (fp)	Frequência Acumulada (FA)	Frequência Percentual Acumulada (FP)
6,27	┊	7,62	3	0,10	10%	3	10%
7,62	┊	8,97	7	0,23	23%	10	33%
8,97	┊	10,32	10	0,33	33%	20	67%
10,32	┊	11,67	6	0,20	20%	26	87%
11,67	┊	13,02	4	0,13	13%	30	100%
			30	1,00	100%		

Representação gráfica: histograma e pólígono de frequências.

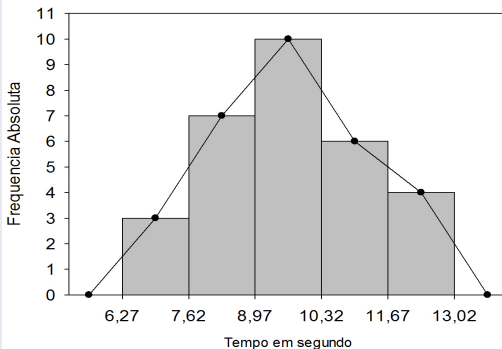


Figura: Histograma e Polígono de frequências relativa ao tempo em segundos para carga de um aplicativo num sistema compartilhado.

Representação gráfica: ogiva.

- O primeiro ponto da ogiva é formado pelo limite inferior da primeira classe e o valor zero, indicando que abaixo do limite inferior da primeira classe não existem observações. Daí por diante, são usados os limites superiores das classes e suas respectivas frequências acumuladas, até a última classe, que acumula todas as observações. Assim, uma ogiva deve começar no valor zero e, se for construída com as frequências relativas acumuladas, terminar com o valor 100.

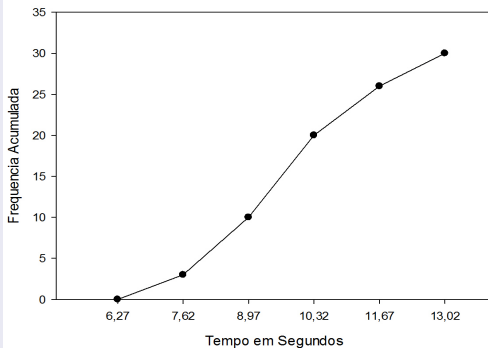


Figura: Ogiva para o tempo em segundos para carga de um aplicativo num sistema compartilhado.