

Apostila de Estatística

Volume 1 - Edição 2007

Curso: Psicologia

Amostragem, Séries Estatísticas, Distribuição de
Frequência, Média, Mediana, Quartil, Percentil e
Desvio Padrão

Prof. Dr. Celso Eduardo Tuna

Estatística

Capítulo 1 - Introdução

1.1 Histórico

A estatística é um ramo da matemática aplicada.

A partir do século XVI começaram a surgir as primeiras análises sistemáticas de registros diversos como os de nascimento, óbitos, riquezas, casamentos. Esses registros eram utilizados para principalmente cobrar impostos.

No século XVIII, Godofredo Achenwall batizou esses estudos como uma nova ciência com o nome de Estatística. Surgiram tabelas mais complexas, representações gráficas e cálculo de probabilidade. Formou-se a ferramenta que através da observação de partes (amostras) chega-se a conclusões sobre um todo (população).

1.2 Método Estatístico

Método é o conjunto de procedimentos dispostos ordenadamente para se chegar a um desejado fim.

Dos métodos científicos pode-se destacar:

Método Experimental: consiste em manter constantes todas as causas (fatores, componentes, variáveis), menos uma, e variar essa última para descobrir seus efeitos, caso existam.

Método Estatístico: diante da impossibilidade de manter as causas constantes, registram-se os resultados dessas variações procurando determinar a influência (os efeitos) de cada uma delas.

1.3 Estatística

A Estatística é parte da Matemática Aplicada que fornece métodos de coleta, organização, descrição, análise e interpretação de dados, úteis nas tomadas de decisão.

Estatística Descritiva: coleta, organização e descrição dos dados.

Estatística Indutiva ou Inferencial: análise e interpretação dos dados. Permite obter conclusões que transcendam os dados obtidos inicialmente, objetivo essencial da Estatística.

Probabilidade: útil para analisar situações que envolvem o acaso. Ex: a decisão de parar de imunizar pessoas com mais de vinte anos contra determinada doença.

1.4 Método Estatístico (Pesquisa)

Exemplos:

- Indústrias realizam pesquisa entre os consumidores para o lançamento de um novo produto
- As pesquisas eleitorais fornecem elementos para que os candidatos direcionem a campanha
- Emissoras de tevê utilizam pesquisas que mostram a preferência dos espectadores para organizar sua programação
- A pesquisa do desempenho dos atletas ou das equipes em uma partida ou em um campeonato interfere no planejamento dos treinamentos

A pesquisa é composta basicamente de 5 fases

1ª Coleta de Dados

Após planejamento e determinação das características mensuráveis do objeto em estudo, inicia-se a coleta de dados. Esta pode ser direta ou indireta.

A coleta direta é feita sobre registros diversos: nascimento, casamento, óbitos, importação, registros escolares; ou ainda quando os dados são coletados diretamente pelo pesquisador através de questionários (ex: censo).

A coleta direta pode ser: contínua; periódica (censos); ocasional

A coleta indireta é uma coleta feita sobre dados colhidos de uma coleta direta (ex: mortalidade infantil)

2ª Crítica dos Dados

Os dados coletados devem ser observados, à procura de falhas e imperfeições, a fim de não causarem erro nos resultados.

Exemplo 1 : Perguntas tendenciosas. Foi realizada a seguinte pesquisa:

O tráfego contribui em maior ou menor grau do que a indústria para a poluição atmosférica ?

Resposta: 45 % para o tráfego e 32 % para a indústria.

A indústria contribui em maior ou menor grau do que o tráfego para a poluição atmosférica ?

Resposta: 24 % para o tráfego e 57 % para a indústria.

Exemplo 2: Preservação da auto-imagem. Em uma pesquisa telefônica 94 % dos entrevistados disseram que lavam as suas mãos após usar o banheiro, mas a observação em banheiros públicos esse percentual cai para 68 %.

Exemplo 3: Más Amostras. As pessoas devem ser escolhidas aleatoriamente para a pesquisa, como por exemplo, numa pesquisa de opinião na rua, deve-se entrevistar somente quem pisou em uma determinada marca pré-determinada na calçada.

Exemplo 4. Más perguntas. A pergunta deve conter o linguajar próprio do entrevistado. Geralmente, se o entrevistado não entender a pergunta, ele responderá qualquer coisa, pois tem vergonha de perguntar.

3ª Apuração dos Dados

É o processamento dos dados obtidos

4ª Exposição dos Dados

Através de tabelas ou gráficos, tornando mais fácil seu exame e aplicação de um cálculo estatístico

5ª Análise dos Resultados

Através de métodos de estatística indutiva ou inferencial obtêm-se conclusões e previsões de um todo através do exame de apenas uma parte desse todo.

Capítulo 2 - População e Amostra

2.1 Variável

Variável é o conjunto de resultados possíveis de um fenômeno. A variável pode ser qualitativa, quando seus valores são expressos por atributos (ex: sexo, cor), ou pode ser quantitativa, quando seus valores são expressos em números.

A variável quantitativa pode ser contínua, quando assume qualquer valor entre dois limites (ex: peso, altura, medições), ou pode ser discreta, quando só pode assumir valores pertencentes a um conjunto enumerável (ex: número de filhos, contagens em geral, números inteiros).

2.2 Precisão

A precisão da medida será automaticamente indicada pelo número de decimais com que se escrevem os valores da variável. Ex: 1,80 m indica uma medição com precisão de centésimos.

2.3 Arredondamento

De acordo com resolução do IBGE

Quando o primeiro algarismo a ser abandonado é 0, 1, 2, 3, ou 4, fica inalterado o último algarismo a permanecer. Ex: 53,24 passa a 53,2 ; 17,3452 passa a 17,3

Quando o primeiro algarismo a ser abandonado é 6, 7, 8, ou 9, aumenta-se de uma unidade o último algarismo a permanecer. Ex: 42,87 passa a 42,9 ; 25,08 passa a 25,1; 53,99 passa a 54,0

Quando o primeiro algarismo a ser abandonado é 5, há duas soluções:

a) Se ao 5 seguir em qualquer casa um algarismo diferente de 0, aumenta-se de uma unidade o último algarismo a permanecer. Ex: 2,352 passa a 2,4 ; 25,6501 passa a 25,7.

b) Se o 5 for o último algarismo ou se ao 5 só se seguirem zeros, o último algarismo a ser conservado só será aumentado de uma unidade se for ímpar. Ex: 24,75 passa a 24,8 ; 24,65 passa a 24,6 ; 24,7500 passa a 24,8 ; 24,6500 passa a 24,6

Exercícios. Arredonde deixando número inteiro:

2,38 = 24,65 = 0,351 = 4,24 = 328,35 =

2,97 = 6,829 = 5,55 = 89,99 =

Exercícios. Arredonde deixando uma casa decimal:

2,38 = 24,65 = 0,351 = 4,24 = 328,35 =

2,97 = 6,829 = 5,55 = 89,99 =

2.4 População e Amostra

População é o conjunto de portadores de, pelo menos, uma característica comum.

Amostra é um subconjunto finito de uma população.

A amostra é escolhida através de processos adequados que garantam o acaso na escolha

2.5 Amostragem

É o processo de colher amostras. Nesse processo, cada elemento da população passa a ter a mesma chance de ser escolhido. Dentre os processos de amostragem pode-se destacar três: amostragem casual ou aleatória simples, amostragem proporcional estratificada e amostragem sistemática.

a) Amostragem casual ou aleatória simples:

É um sorteio, por exemplo, para retirar uma amostra de 9 alunos de uma sala de 90 alunos, utiliza-se um sorteio com todos os números dos alunos escritos em papéis dentro de um saco. Para amostras grandes utiliza-se a Tabela de Números Aleatórios (Página 40). Assim para o exemplo da sala de aula, utilizando dois algarismos, através da leitura da primeira linha (escolhida através de sorteio), obtém-se:

Como a população vai de 1 a 90 escolhe-se os 9 primeiros números dentro dessa faixa:

b) Amostragem proporcional estratificada:

É comum termos populações que se dividam em subpopulações (estratos) e como cada estrato pode ter um comportamento diferente do outro, a amostra deve considerar a existência desses estratos e a sua proporção em relação à população.

Exemplo: supondo que uma sala de aula seja composta de 54 meninos e 36 meninas. Determine uma amostra de 9 pessoas:

Sexo	População	Cálculo Proporcional Regra de três simples	Amostra
Masculino	54	$54 \times 9 / 90 = 5,4$	5
Feminino	36	$36 \times 9 / 90 = 3,6$	4
Total	90	9	9

Posteriormente, utiliza-se a tabela de números aleatórios para escolher 5 meninos e 4 meninas.

Verifica-se que foi realizado um arredondamento dos números 5,4 e 3,6. Esse arredondamento é efetuado utilizando as regras de arredondamento.

Exercício: Em uma escola existem 250 alunos, distribuídos em séries conforme a tabela. Obtenha uma amostra de 40 alunos e preencha a tabela.

Séries	População	Cálculo Proporcional	Amostra
1 ^a	35		
2 ^a	32		
3 ^a	30		
4 ^a	28		
5 ^a	35		
6 ^a	32		
7 ^a	31		
8 ^a	27		
Total	250		40

c) Amostragem sistemática

É quando a amostragem é feita através de um sistema possível de ser aplicado pois a população já se encontra ordenada.

Exemplo 1: em uma linha de produção, a cada 10 itens fabricados, retira-se 1 para inspeção, tem-se uma amostra de 10 % da população.

Exemplo 2: em uma rua com 900 prédios, deseja-se uma amostra de 50. $900/50 = 18$ (50 grupos de 18 prédios cada). Faz-se um sorteio entre 1 e 18, por exemplo 4, então pesquisariamos o 4^o prédio da rua, o 22^o, o 40^o, 58^o, assim por diante.

Exercícios de População e Amostra

1) Uma universidade apresenta o seguinte quadro relativo aos seus alunos do curso de Matemática. Obtenha uma amostra proporcional estartificada de 100 alunos.

Série	Qtde	Amostra
1 ^a	85	
2 ^a	70	
3 ^a	80	
4 ^a	75	
Total		100

2) Uma cidade X apresenta o seguinte quadro relativo às suas escolas de 1º grau:

Escola	Homens	Mulheres	Total	Amostra		
				Homens	Mulheres	Total
A	80	95				
B	102	120				
C	110	92				
D	134	228				
E	150	130				
F	300	290				
Total						120

Obtenha uma amostra proporcional estratificada de 120 estudantes

3) Utilizando a tabela de números aleatórios, obtenha uma amostra de 10 pessoas de uma sala de aula com 85 alunos, utilize a 10^a e a 11^a coluna para começar o sorteio.

4) Ordene uma amostra de 15 elementos de uma população ordenada formada por 210 elementos, sabendo que o elemento de ordem 149 a ela pertence ?

Capítulo 3 - Séries Estatísticas

3.1 Séries Estatísticas

Série estatística é toda tabela que apresenta a distribuição de um conjunto de dados estatísticos em função da época, do local, ou da espécie.

Pode-se classificar em: histórica, geográfica, específica

a) Séries históricas (cronológicas, temporais) - descrevem os valores da variável, em determinado local, em função do tempo

Exemplo: **Tabela – Analfabetismo na faixa de 15 anos ou mais - Brasil - 1900/2000**

População de 15 anos ou mais

Ano	Total(1)	Analfabeta(1)	Taxa de Analfabetismo
1900	9.728	6.348	65,3
1920	17.564	11.409	65,0
1940	23.648	13.269	56,1
1950	30.188	15.272	50,6
1960	40.233	15.964	39,7
1970	53.633	18.100	33,7
1980	74.600	19.356	25,9
1991	94.891	18.682	19,7
2000	119.533	16.295	13,6

Fonte: IBGE, Censo Demográfico.

Nota: (1) Em milhares

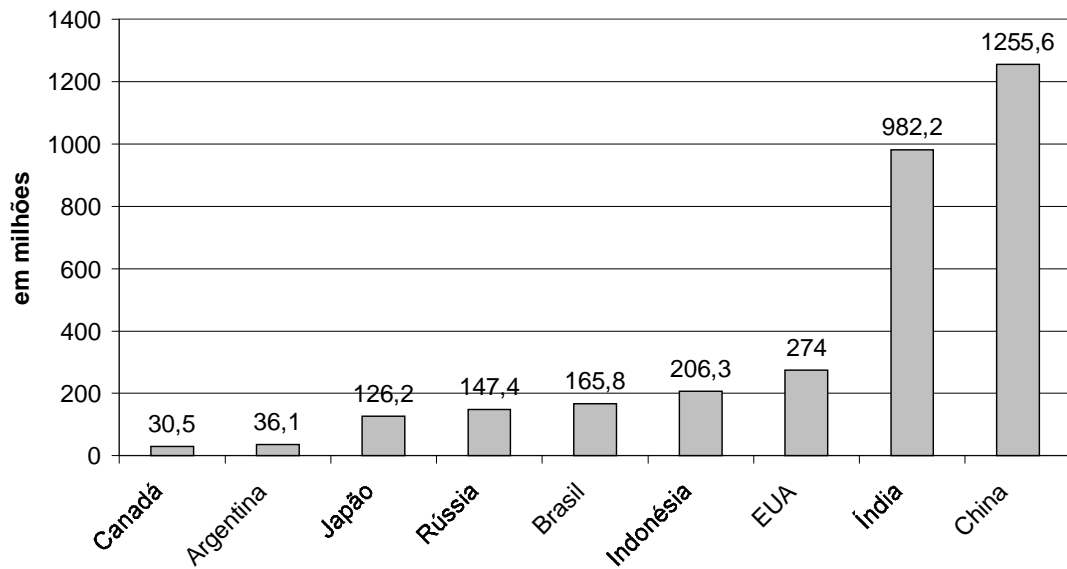
b) Séries geográficas (espaciais, territoriais ou de localização) - descrevem os valores da variável, em um determinado instante, em função da região

Exemplo:

População Mundial	
Em milhões de pessoas - 1998	
Canadá	30,5
Argentina	36,1
Japão	126,2
Rússia	147,4
Brasil	165,8
Indonésia	206,3
EUA	274
Índia	982,2
China	1255,6

Fonte: O Estado de São Paulo, 01/01/2000

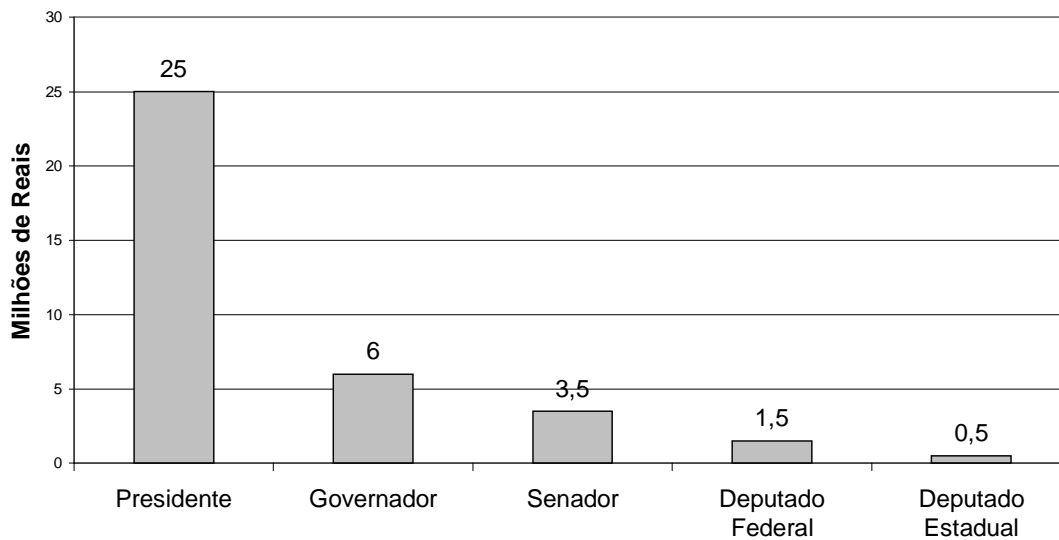
População Mundial em 1998



c) Séries Específicas (categóricas) - descrevem os valores da variável, em um determinado instante e local, segundo especificações.

Custo médio das campanhas eleitorais em 1998, segundo estimativa dos candidatos em milhões de reais. Fonte: TSE	
Presidente	25
Governador	6
Senador	3,5
Deputado Federal	1,5
Deputado Estadual	0,5

Custo médio das campanhas eleitorais em 1998, segundo estimativa dos candidatos em milhões de reais. Fonte: TSE



d) Séries Conjugadas - Tabela de Dupla Entrada

É a união de duas séries em uma só tabela

Exemplo:

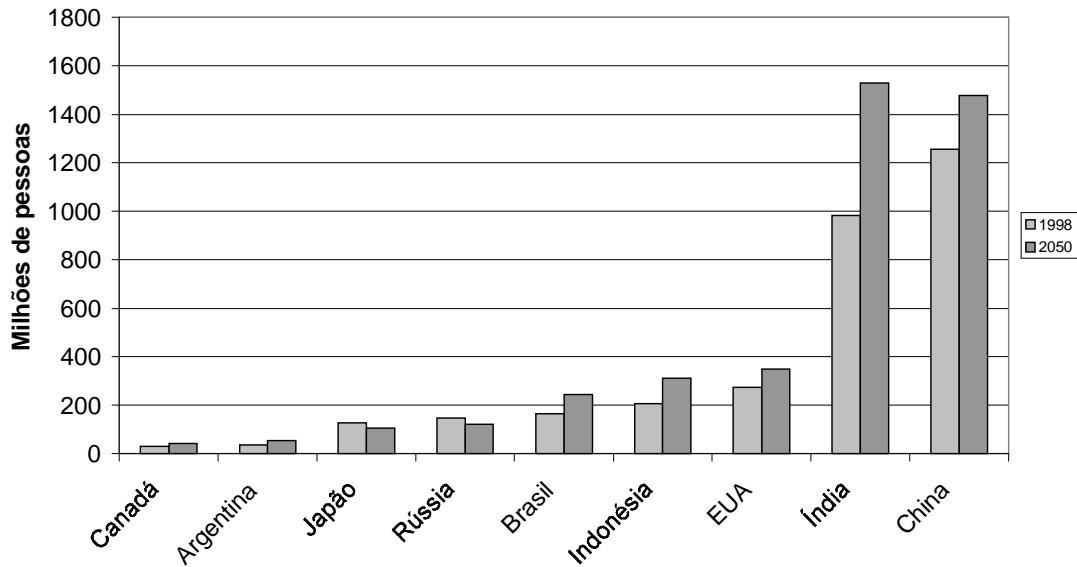
População Mundial - em milhões de pessoas		
País	1998	2050
Canadá	30,5	42,3
Argentina	36,1	54,5
Japão	126,2	104,9
Rússia	147,4	121,2
Brasil	165,8	244,2
Indonésia	206,3	311,8
EUA	274	349,3
Índia	982,2	1528,8
China	1255,6	1477,7

Fonte: O Estado de São Paulo, 01/01/2000

O exemplo acima é uma série geográfica-histórica

Podem também existir séries conjugadas de três ou mais entradas, fato mais raro, pois dificulta a interpretação dos dados.

População Mundial - em milhões de pessoas



3.2 - Distribuição de frequência

Será tratado em capítulo a parte devido a sua importância.

Exemplo:

Idade na morte causada por arma de fogo	
Idade na Morte	Frequência
15 25	22
25 35	10
35 45	6
45 55	2
55 65	4
65 75	5
75 85	1

3.3 Dados Absolutos e Dados Relativos

Dados Absolutos: são resultantes de uma coleta direta, sem outra manipulação senão a contagem

Dados Relativos: são resultantes de comparações, há um tratamento matemático dos dados para uma melhor interpretação.

3.3.1 - As porcentagens

a) Considere a série:

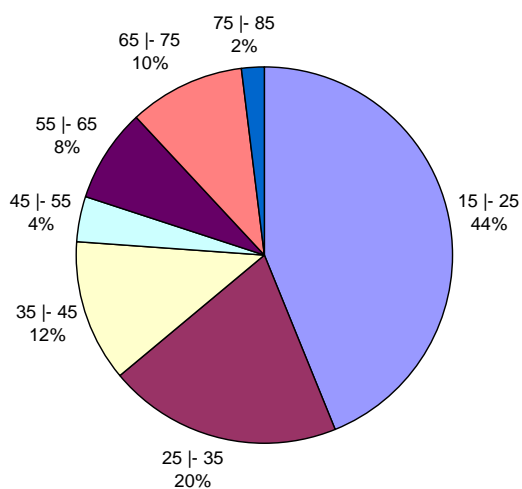
Idade na morte causada por arma de fogo	
Idade na Morte	Frequência
15 - 25	22
25 - 35	10
35 - 45	6
45 - 55	2
55 - 65	4
65 - 75	5
75 - 85	1

Calculando a porcentagem das pessoas em cada faixa etária, pode-se preencher uma nova coluna

Idade na Morte	Frequência	%
15 - 25	22	44
25 - 35	10	20
35 - 45	6	12
45 - 55	2	4
55 - 65	4	8
65 - 75	5	10
75 - 85	1	2
Total	50	100

Pode-se agora tirar uma melhor conclusão e também construir um gráfico de setores (pizza).

Idade da Morte causada por arma de fogo



3.3.2 - Os índices

Os índices são razões entre duas grandezas independentes. Ex:

Relação candidato vaga = Qtde de candidatos / Qtde de vagas

Densidade demográfica = população / área de uma superfície

Renda per capita = renda total de uma população / população

3.3.3 - Os Coeficientes

Os coeficientes são razões entre o número de ocorrências e o número total. É a porcentagem expressa na forma unitária. Ex:

Coeficiente de evasão escolar = n° de alunos evadidos / n° inicial de alunos

Coeficiente de aproveitamento escolar = n° de alunos aprovados / n° final de alunos

3.3.4 - As Taxas

As taxas são os coeficientes multiplicados por uma potência de 10, 100, 1000, etc para tornar o resultado mais inteligível (claro) Ex:

Taxas de mortalidade = coeficiente de mortalidade x 1000 (lê-se mortes a cada 1000 habitantes)

Taxa de evasão escolar = coeficiente de evasão escolar x 100

Exercícios:

Exercício 1 - Considere a tabela abaixo:

Ano	Qtde de Analfabetos no Brasil acima de 15 anos em milhares de hab.	% de aumento
1960	40233	_____
1970	53633	
1980	74600	
1991	94891	
2000	119533	

Complete a tabela com uma coluna de porcentagem de aumento de um período para o outro. Não utilize casas decimais, apenas números inteiros.

Exercício 2 - Considerando que o Brasil, em 2000, apresentou:

População: 164 milhões de habitantes

Superfície: 8 511 996 km²

Nascimentos: 6,2 milhões

Óbitos: 3,8 milhões

Calcule:

a) o índice de densidade demográfica

b) a taxa de natalidade

c) a taxa de mortalidade

Exercício 3 - Em certa eleição municipal foram obtidos os seguintes resultados

Candidato	% do total de votos	Número de votos
A	26	
B	24	
C	22	
Branco e nulos		196

Determine o número de votos obtido pelo candidato vencedor.

Exercício 4 : A tabela abaixo apresenta a variação percentual das vendas industriais de aparelhos domésticos, comparando o período de julho e agosto de 2003 com o período de julho e agosto de 2004.

Vendas industriais de aparelhos domésticos Variação percentual jul/ago 2003 e jul/ago 2004	
Refrigeradores	15,06
Freezers verticais	4,97
Freezers horizontais	42,61
Lavadoras automáticas	- 18,18
Fogões	- 0,17
Condicionadores de ar	83,45

Supondo que no período de jul/ago de 2003 tenham sido vendidas 200.000 lavadoras automáticas, determine o número de unidades vendidas no mesmo período de 2004.

Capítulo 4 - Distribuição de Frequência

4.1 Tabela Primitiva e Rol

Tabela primitiva - elementos da variável ainda não foram numericamente organizados

Ex:

Total de pontos (acertos) obtidos por 40 alunos em um teste de 175 questões

166	160	161	150	162	160	165	167	164	160
162	161	168	163	156	173	160	155	164	168
155	152	163	160	155	155	169	151	170	164
154	161	156	172	153	157	156	158	158	161

Rol - é a tabela primitiva ordenada (crescente ou decrescente).

Ex:

150	154	155	157	160	161	162	164	166	169
151	155	156	158	160	161	162	164	167	170
152	155	156	158	160	161	163	164	168	172
153	155	156	160	160	161	163	165	168	173

4.2 Distribuição de frequência

Com isso pode-se construir uma tabela denominada Distribuição de Frequência, sendo a frequência o número de elementos relacionados a um determinado valor da variável.

Ex:

Pontos	Frequência	Pontos	Frequência	Pontos	Frequência
150	1	158	2	167	1
151	1	160	5	168	2
152	1	161	4	169	1
153	1	162	2	170	1
154	1	163	2	172	1
155	4	164	3	173	1
156	3	165	1		
157	1	166	1	total	40

Para uma melhor visualização e economia de espaço, agrupam-se os valores em intervalos de classe.

Ex:

Total de pontos (acertos) obtidos em um teste de 175 questões por 40 alunos	
Total de pontos	Frequência
150 - 154	4
154 - 158	9
158 - 162	11
162 - 166	8
166 - 170	5
170 - 174	3
Total	40

Para a confecção dessa tabela pode-se pular o passo anterior, ou seja, do rol já partir para a tabela de distribuição de frequências com intervalos de classe.

4.3 Elementos de uma distribuição de frequência

a) **Classes de frequência:** são os intervalos de variação da variável, representados por i , sendo $i = 1, 2, 3, 4, \dots, k$, onde k é o número total de classes.

Em nosso exemplo $k = 6$

b) **Limites da classe:** são os extremos de cada classe.

Limite superior L_i Limite inferior l_i

O símbolo $l_i | - L_i$ significa inclusão de l_i e exclusão de L_i

$$l_2 = 154 \text{ e } L_2 = 158$$

c) **Amplitude de um intervalo de classe (h)** é a medida do intervalo que define a classe

$$h = L_i - l_i \quad h_2 = 154 - 158 = 4$$

d) **Amplitude total da distribuição (AT)** é a diferença entre o limite superior da última classe (limite superior máximo) e o limite inferior da primeira (limite inferior mínimo).

$$AT = L(\text{max}) - l(\text{min})$$

$$AT = 174 - 150 = 24$$

$$\text{Deve-se notar que } AT/h = k \quad 24/4 = 6$$

e) **Amplitude amostral (AA)** : é a diferença entre o valor máximo e o valor mínimo da amostra

$$AA = x(\text{máx}) - x(\text{mín}) \quad AA = 173 - 150 = 23$$

f) **Ponto médio** de uma classe (x_i) : é o ponto que divide o intervalo de classe em duas partes iguais

$$x_i = (l_i + L_i)/2 \quad x_2 = (154 + 158)/2 = 156$$

f) **Frequência simples ou absoluta:** é o número de observações correspondentes a essa classe ou a esse valor

$$f_1 = 4 \quad f_2 = 9 \quad f_3 = 11 \quad f_4 = 8 \quad f_5 = 5 \quad f_6 = 3$$

$$\sum_{i=1}^k f_i = n \quad \sum_{i=1}^6 f_i = 40$$

4.4 Número de Classes, Intervalos de Classe

Determinação do número de classes: utiliza-se a regra de Sturges (obs: não é obrigatório, é apenas uma orientação)

$k = 1 + 3,3 \cdot \log n$ onde, k é o número de classes e n é o número total de dados. Esta fórmula nos permite obter a seguinte tabela

n	k
3 - 5	3
6 - 11	4
12 - 22	5
23 - 46	6
47 - 90	7
91 - 181	8
182 - 362	9

Para determinação do intervalo de classe h aplica-se

$$h = \frac{AA}{k}$$

Quando o resultado não é exato, deve-se arredondá-lo para mais.

No caso $h = \frac{173 - 150}{6} = 3,8 = 4$, ou seja, 6 classes de intervalo 4.

Exercício: As notas obtidas por 50 alunos de uma classe foram:

1	2	3	4	5	6	6	7	7	8
2	3	3	4	5	6	6	7	8	8
2	3	4	4	5	6	6	7	8	9
2	3	4	5	5	6	6	7	8	9
2	3	4	5	5	6	7	7	8	9

Complete a distribuição de freqüência abaixo

i	Notas	x_i	f_i
	0 - 2		
	2 - 4		
	4 - 6		
	6 - 8		
	8 - 10		
	Total		50

4.5 Tipos de freqüências

- a) Freqüência Simples ou Absoluta (f_i) : é o valor que representa o número de dados de uma classe, onde :

$$\sum_{i=1}^k f_i = n$$

b) Frequência Relativa (fr_i): é a porcentagem entre a frequência simples e a frequência total:

$$fr_i = \frac{f_i}{\sum_{i=1}^k f_i} \cdot 100[\%]$$

No exemplo: $fr_3 = 11/40 = 0,275 \times 100 = 27,5 \%$

É obvio que: $\sum_{i=1}^k fr_i = 100\%$

O propósito das frequências relativas é o de permitir a análise e facilitar comparações.

c) Frequência Acumulada (F_i): é o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe.

$$F_k = f_1 + f_2 + f_3 + \dots + f_k \quad \text{ou} \quad F_k = \sum_{i=1}^k f_i$$

No exemplo $F_3 = f_1 + f_2 + f_3 = 4+9+11=24$, o que significa que existem 24 alunos com estatura inferior a 162 cm (limite superior do intervalo da terceira classe)

d) Frequência Acumulada relativa (Fr_i): é a porcentagem entre a frequência relativa acumulada da classe e a frequência total da distribuição.

$$Fr_i = \frac{F_i}{\sum_{i=1}^k f_i} \cdot 100[\%]$$

No exemplo temos $Fr_3 = 24/40 = 0,6 = 60 \%$, o que significa que 60 % dos alunos acertaram menos de 162 questões

Pode-se então montar a seguinte tabela:

i	Total de Pontos	x_i	f_i	fr_i (%)	F_i	Fr_i (%)
1	150 - 154	152	4	10,00	4	10,00
2	154 - 158	156	9	22,50	13	32,50
3	158 - 162	160	11	27,50	24	60,00
4	162 - 166	164	8	20,00	32	80,00
5	166 - 170	168	5	12,50	37	92,50
6	170 - 174	172	3	7,50	40	100,00
	Total		40	1,000		

Que nos ajuda a responder:

- 1) Quantos alunos acertaram entre 154, inclusive, e 158 questões ? Resp. 9 alunos
- 2) Qual a porcentagem de alunos com total de pontos inferior a 154? Resp. 10%
- 3) Quantos alunos acertaram menos que 162 questões ? Resp. 24 alunos
- 4) Quantos alunos obtiveram um total de pontos não inferior a 158? Resp. $40-13 = 27$ alunos

4.6 Distribuição de Freqüência sem Intervalo de Classe

Quando se trata de variável discreta de variação relativamente pequena, cada valor pode ser tomado como um intervalo de classe, tomando a seguinte forma:

Os resultados de um lançamento de um dado 50 vezes foram os seguintes:

6	5	2	6	4	3	6	2	6	5
1	6	3	3	5	1	3	6	3	4
5	4	3	1	3	5	4	4	2	6
2	2	5	2	5	1	3	6	5	1
5	6	2	4	6	1	5	2	4	3

i	resultados	f_i	fr_i	F_i	Fr_i
1	1				
2	2				
3	3				
4	4				
5	5				
6	6				
	Total	50	1,000		

Exercício: Complete a tabela abaixo e responda:

i	Horas de estudo por semana	x_i	f_i	fr_i	F_i	Fr_i
1	0 - 5		5			
2	5 - 10		96			
3	10 - 15		57			
4	15 - 20		25			
5	20 - 25		11			
6	25 - 30		6			
	Total			1,000		

Qual a porcentagem de pessoas que estudam menos de 15 horas ?

Qual a porcentagem de pessoas que estudam 20 ou mais horas ?

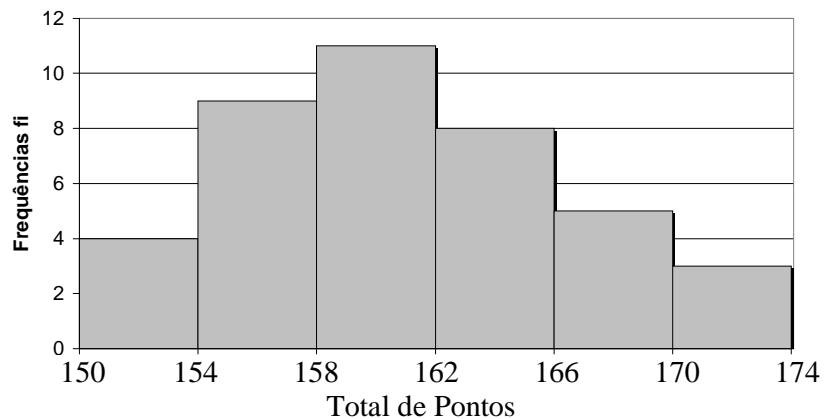
4.7 Representação Gráfica de uma Distribuição de Frequência

Pode-se ser representado basicamente por um histograma, por um polígono de frequência ou por um polígono de frequência acumulada.

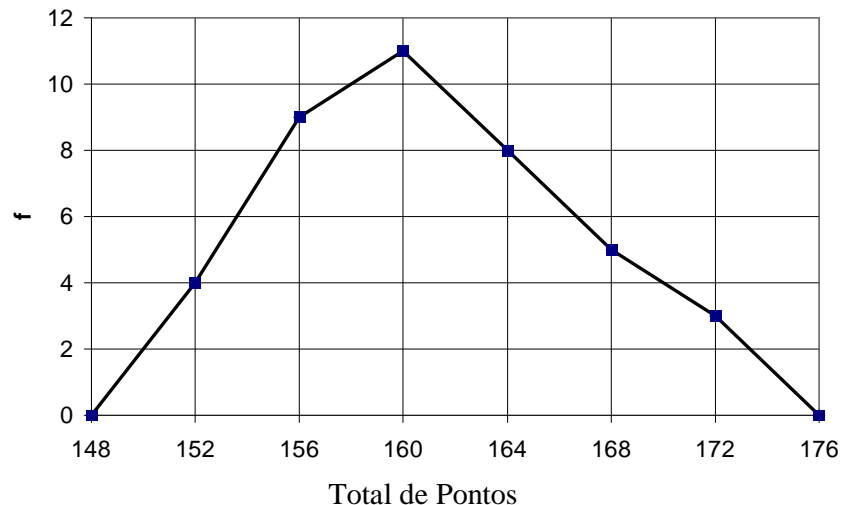
- a) **Histograma:** O histograma é formado por um conjunto de retângulos justapostos, cujas bases se localizam sobre o eixo horizontal, de tal modo que seus pontos médios coincidam com os pontos médios dos intervalos de classe. Seja o exemplo:

i	Total de Pontos	x_i	f_i	F_i
1	150 - 154	152	4	4
2	154 - 158	156	9	13
3	158 - 162	160	11	24
4	162 - 166	164	8	32
5	166 - 170	168	5	37
6	170 - 174	172	3	40
	Total		40	

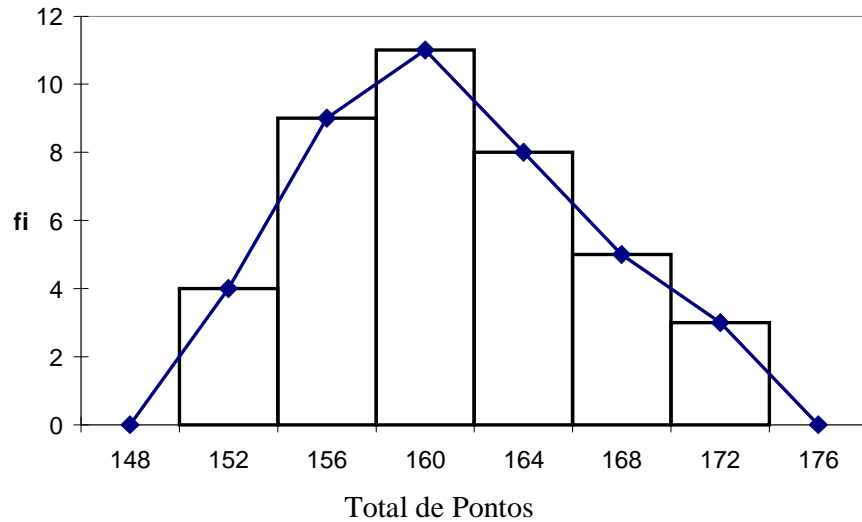
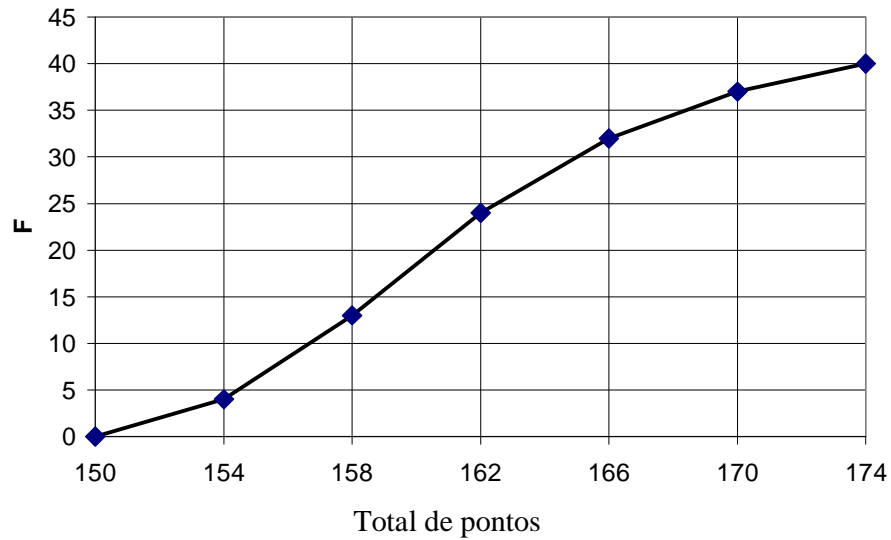
Histograma



- b) **Polígono de frequência:** É um gráfico em linha, sendo as frequências marcadas sobre perpendiculares ao eixo horizontal, levantadas pelos pontos médios dos intervalos de classe.



- c) Polígono de frequência acumulada: É traçado marcando-se as frequências acumuladas sobre perpendiculares ao eixo horizontal, levantadas nos pontos correspondentes aos limites superiores dos intervalos de classe.



Polígono de frequência com o histograma

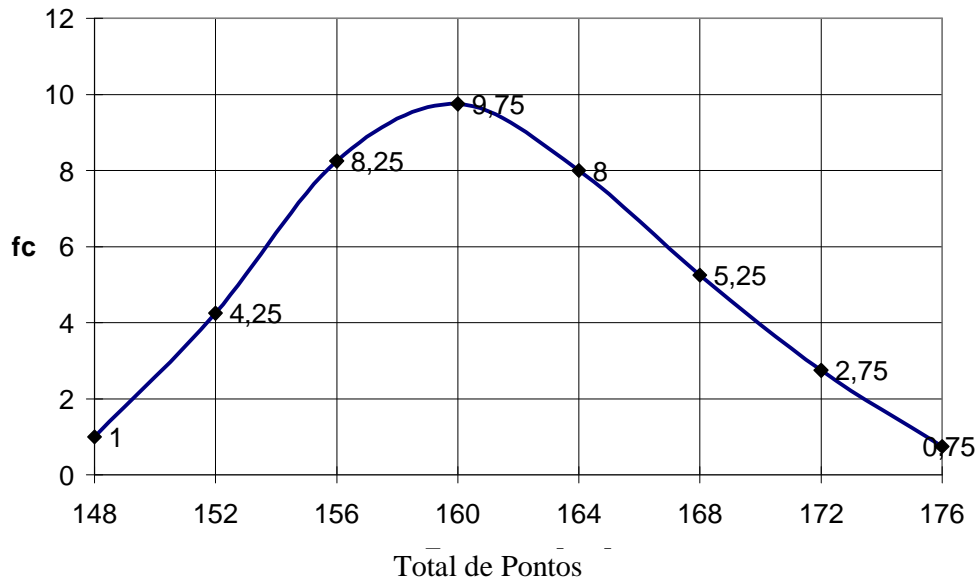
4.8 - A Curva de Frequência. Curva Polida

O polígono de frequência nos fornece uma imagem real e a curva uma imagem tendencial. A curva polida de uma amostra limitada se assemelha mais a curva resultante de um grande número de dados, do que o polígono de frequência obtido da mesma amostra limitada. Utiliza-se uma nova frequência, denominada calculada (f_c).

$$f_{c_i} = \frac{f_{(i-1)} + 2 \cdot f_i + f_{(i+1)}}{4}$$

No exemplo anterior tem-se:

i	Total de Pontos	x_i	f_i	F_i	fc
0	146 - 150	148	0	0	$(0+2*0+4)/4 = 1$
1	150 - 154	152	4	4	$(0+2*4+9)/4 = 4,25$
2	154 - 158	156	9	13	$(4+2*9+11)/4 = 8,25$
3	158 - 162	160	11	24	$(9+2*11+8)/4 = 9,75$
4	162 - 166	164	8	32	$(11+2*8+5)/4 = 8$
5	166 - 170	168	5	37	$(8+2*5+3)/4 = 5,25$
6	170 - 174	172	3	40	$(5+2*3+0)/4 = 2,75$
7	174 - 178	176	0	40	$(3+2*0+0)/4 = 0,75$
	Total		40		



Exercício - Construa o histograma, o polígono de frequência, o polígono de frequência acumulada e a curva polida da seguinte distribuição.

i	Total de Faltas de uma sala com 60 alunos	x_i	f_i	fc_i	F_i
0					
1	0 - 2		5		
2	2 - 4		15		
3	4 - 6		25		
4	6 - 8		10		
5	8 - 10		5		
6					

Capítulo 5 - Medidas de Posição

5.1 Introdução

Até agora os estudos de **distribuição de frequência** efetuados nos **permite localizar a maior e menor concentração dos valores de uma dada distribuição**. No entanto, para destacar as **tendências características** necessita-se de elementos típicos da distribuição que são as:

- Medidas de posição
- Medidas de variabilidade ou dispersão
- Medidas de assimetria
- Medidas de curtose

As **medidas de posição** nos orienta quanto a posição da distribuição **em relação ao eixo horizontal**. As medidas mais importantes são as medidas de **tendência central** (os dados tendem a se agrupar em torno de valores centrais). Dentre elas destacam-se:

- A média aritmética
- A mediana
- A moda

Outras medidas de posição são as separatrizes que são:

A mediana
Os quartis
Os percentis

5.2 Media Aritmética (\bar{X})

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

onde x_i são os valores da variável e n o número de valores.

a) **Desvio em relação a média (d_i)** $d_i = x_i - \bar{X}$

b) **Propriedades:** $\sum_{i=1}^n d_i = 0$

A soma algébrica dos desvio em relação a média é nula

Somando-se (ou subtraindo-se) uma constante (c) de todos os valores de uma variável, a média do conjunto fica aumentada (ou diminuída) dessa constante.

Multiplicando-se (ou dividindo-se) uma constante (c) de todos os valores de uma variável, a média do conjunto fica multiplicada (ou dividida) por essa constante.

Exemplo: Seja a nota de 10 alunos: 8, 9, 7, 6, 10, 5,5, 5, 6,5, 7,5, 8,5

$$\text{A média é } \bar{x} = \frac{8+9+7+6+10+5,5+5+6,5+7,5+8,5}{10} = 7,3$$

Desvios:

8 - 7,3	0,7
9 - 7,3	1,7
7 - 7,3	-0,3
6 - 7,3	-1,3
10 - 7,3	2,7
5,5 - 7,3	-1,8
5 - 7,3	-2,3
6,5 - 7,3	-0,8
7,5 - 7,3	0,2
8,5 - 7,3	1,2
Total	0,0

c) para dados agrupados (distribuição de freqüência sem intervalos de classe)

Seja a seguinte distribuição:

nº de filhos (x_i) que se deseja ter	f_i	$f_i \cdot x_i$
0	2	0
1	6	6
2	10	20
3	12	36
4	4	16
Total	34	78

$$\bar{x} = \frac{\sum_{i=1}^n (f_i \cdot x_i)}{\sum_{i=1}^n f_i}$$

tem-se então: $\bar{x} = \frac{78}{34} = 2,294 \cong 2,3$

d) para dados agrupados (distribuição de freqüência com intervalos de classe). Adota-se o seguinte: todos os valores incluídos em um determinado intervalo de classe coincidem com o seu ponto médio.

Seja a seguinte distribuição:

i	Total de pontos	x_i	f_i	$f_i \cdot x_i$
1	150 - 154	152	4	608
2	154 - 158	156	9	1404
3	158 - 162	160	11	1760
4	162 - 166	164	8	1312
5	166 - 170	168	5	840
6	170 - 174	172	3	516
	Total		40	6440

tem-se então: $\bar{x} = \frac{6440}{40} = 161$ pontos

Exercício 1 - Complete a tabela e calcule a média aritmética da distribuição.

Qtde de cursos de extensão realizados por ano (x_i) pelos alunos do 3º Mat	f_i	$f_i \cdot x_i$
1	2	
2	4	
3	6	
4	8	
5	3	
6	1	

Exercício 2 - Complete a tabela e calcule a média aritmética da distribuição.

i	Salário Mensal dos alunos do 3º Mat [R\$]	x_i	f_i	$f_i \cdot x_i$
1	450 550		8	
2	550 650		10	
3	650 750		11	
4	750 850		16	
5	850 950		13	
6	950 1050		5	
7	1050 1150		1	
	Total			

e) Processo breve

Há uma mudança de variável x por outra y, tal que:

$$y_i = \frac{x_i - x_0}{h}$$

x_0 é uma constante escolhida convenientemente entre os pontos médios da distribuição, de preferência o de maior valor de frequência, e h é o intervalo de classe. A média então é calculada por:

$$\bar{x} = x_0 + \frac{\left(\sum_{i=1}^n (f_i \cdot y_i) \right) \cdot h}{\sum_{i=1}^n f_i}$$

Exemplo: Escolhendo $x_0 = 160$ e como $h = 4$

i	Total de Pontos	x_i	f_i	y_i	$f_i \cdot y_i$
1	150 - 154	152	4	-2	-8
2	154 - 158	156	9	-1	-9
3	158 - 162	160	11	0	0
4	162 - 166	164	8	1	8
5	166 - 170	168	5	2	10
6	170 - 174	172	3	3	9
	Total		40		10

Então: $\bar{x} = 160 + \frac{10 \cdot 4}{40} = 161$ pontos

Exercício 3: Pelo processo breve, calcule a média aritmética da distribuição.

i	Salário Mensal dos alunos do 3º Mat [R\$]	x_i	f_i	y_i	$f_i \cdot y_i$
1	450 - 550		8		
2	550 - 650		10		
3	650 - 750		11		
4	750 - 850		16		
5	850 - 950		13		
6	950 - 1050		5		
7	1050 - 1150		1		
	Total				

Exercício 4: Pelo processo breve, calcule a média aritmética da distribuição.

i	Valor da hora aula de profissionais da educação [R\$]	x_i	f_i	y_i	$f_i \cdot y_i$
1	30 50		2		
2	50 70		8		
3	70 90		12		
4	90 110		10		
5	110 130		5		
	Total				

5.3 A Moda (Mo)

Denomina-se moda o valor que ocorre com maior frequência em uma série de valores.

Caso 1) Dados não agrupados.

Basta procurar o valor que mais se repete. Ex:

3,4,5,6,6,6,6,7,7,8,9 A série tem moda igual a 6 (valor modal 6)

Pode acontecer também uma série sem valor modal. Ex:

1,2,3,4,5,6,7,8,9 série amodal

Pode acontecer também uma série com mais de uma moda. Ex:

1,2,2,2,3,4,5,6,6,6,7,8,9 a série tem duas modas (2 e 6) - série bimodal

Caso 2) Dados agrupados.

a) **sem intervalos de classe.** Basta identificar o valor da variável que possui maior frequência. Ex:

Seja a seguinte distribuição: $Mo = 3$

nº de filhos (x_i) que se deseja ter	f_i
0	2
1	6
2	10
3	12
4	4
Total	34

b) com intervalos de classe. A classe com maior frequência é denominada classe modal, o cálculo da moda bruta é semelhante ao do ponto médio do intervalo de classe.

$$Mo = x_i = \frac{\ell + L}{2}$$

Ex: Seja a distribuição:

i	Total de pontos	x_i	f_i
1	150 - 154	152	4
2	154 - 158	156	9
3	158 - 162	160	11
4	162 - 166	164	8
5	166 - 170	168	5
6	170 - 174	172	3
	Total		40

Então: a classe modal é $i = 3$, logo $Mo = 160$ pontos

Exercício: Calcule a moda da seguinte distribuição:

i	Salário Mensal dos alunos do 3º Mat [R\$]	f_i
1	450 - 550	8
2	550 - 650	10
3	650 - 750	11
4	750 - 850	16
5	850 - 950	13
6	950 - 1050	5
7	1050 - 1150	1
	Total	64

5.4 Mediana (Md)

A mediana é o número que se encontra no centro de uma série de números, ou seja, separa os valores em dois subconjuntos de mesmo número de elementos.

Caso 1) Dados não agrupados

Dada uma série de valores:

5,13,10,2,18,15,6,16,9

Deve-se então ordená-los:

2,5,6,9,10,13,15,16,18

Determina-se então o valor central que é 10 (4 valores para cada lado) $Md = 10$

Se a série tiver número par de valores, a mediana é a média dos dois valores centrais:

2,5,6,9,10,15,16,18 $Md = (9+10)/2 = 9,5$

Caso 2) Dados agrupados

No caso de distribuição de frequência deve-se primeiramente determinar a frequência acumulada. Determina-se então, o valor que divide a distribuição em duas partes iguais. Aplica-se então:

$$\frac{\sum f_i}{2}$$

a) **sem intervalos de classe.** Dada a série:

nº de filhos (x_i) que se deseja ter	f_i	F_i
0	2	2
1	6	8
2	10	18
3	12	30
4	4	34
Total	34	

Então: $\frac{\sum f_i}{2} = \frac{34}{2} = 17$

A menor frequência acumulada que supera esse valor é 18, que corresponde ao valor 2 da variável.

$$Md = 2$$

No caso de $\frac{\sum f_i}{2} = F_i$ acontecer, a mediana será dada por: $Md = \frac{x_i + x_{i+1}}{2}$. Exemplo:

i	nº de filhos (x_i) que se deseja ter	f_i	F_i
1	0	2	2
2	1	6	8
3	2	10	18
4	3	12	30
5	4	6	36
	Total	36	

$\frac{\sum f_i}{2} = 18 = F_3$, então: $Md = \frac{2+3}{2} = 2,5$

Exercícios:

1) Calcule a mediana das seguintes distribuições:

i	Qtde de anos de estudo (x_i)	f_i	F_i
1	13	6	
2	14	14	
3	15	24	
4	16	16	
5	17	8	
	Total		

i	Qtde de disciplinas em dependência	f _i	F _i
1	0	2	
2	1	5	
3	2	9	
4	3	7	
5	4	6	
6	5	3	
	Total		

b) **com intervalos de classe:** segue-se os seguintes passos:

1º - Determina-se as freqüências acumuladas

2º - Calcula-se $\frac{\sum f_i}{2}$

3º - Marca-se a classe correspondente a freqüência acumulada imediatamente superior a $\frac{\sum f_i}{2}$

(classe mediana) e emprega-se a fórmula:

$$Md = \ell_i + \frac{\left[\frac{\sum f_i}{2} - F(\text{ant}) \right] \cdot h}{f_i}$$

onde: ℓ é o limite inferior da classe mediana

F(ant) é a freqüência acumulada da classe anterior a classe mediana

h é a amplitude do intervalo da classe mediana

f_i é a freqüência do intervalo da classe mediana

Exemplo:

i	Total de pontos	f _i	F _i
1	150 - 154	4	4
2	154 - 158	9	13
3	158 - 162	11	24
4	162 - 166	8	32
5	166 - 170	5	37
6	170 - 174	3	40
	Total	40	

$$\frac{\sum f_i}{2} = \frac{40}{2} = 20, \text{ logo classe mediana é } i = 3 \quad \ell = 158 \quad F(\text{ant}) = 13 \quad h = 4 \quad f_3 = 11$$

$$Md = 158 + \frac{[20 - 13] \cdot 4}{11} = 158 + 2,5 = 160,5$$

No caso de $\frac{\sum f_i}{2} = F_i$ acontecer, a mediana será o limite superior da classe correspondente.

Exercício: Calcule a mediana das seguintes distribuições:

i	Salário Mensal dos alunos do 3º Mat [R\$]	f _i	F _i
1	450 - 550	8	
2	550 - 650	10	
3	650 - 750	11	
4	750 - 850	16	
5	850 - 950	13	
6	950 - 1050	5	
7	1050 - 1150	1	
	Total	64	

i	Valor da hora aula de profissionais da educação [R\$]	f _i	F _i
1	30 - 50	2	
2	50 - 70	8	
3	70 - 90	12	
4	90 - 110	10	
5	110 - 130	5	
	Total		

5.5 Os Quartis

Denomina-se quartis os valores de uma série que a dividem em quatro partes iguais. Portanto, há três quartis. São mais aplicados em distribuição de frequência com intervalos de classe.

Primeiro Quartil (Q₁) - 25 % dos dados são menores que ele e os 75 % restantes são maiores.

Segundo Quartil (Q₂) - coincide com a mediana, 50 % para cada lado.

Terceiro Quartil (Q₃) - 75 % dos dados são menores que ele e os 25 % restantes são maiores.

Para o caso de dados agrupados, basta aplicar: $\frac{k \sum f_i}{4}$, sendo k o número de ordem do quartil.

Então:

$Q_1 = l_i + \frac{\left[\frac{\sum f_i}{4} - F(\text{ant}) \right] \cdot h}{f_i}$	$Q_2 = l_i + \frac{\left[\frac{2 \cdot \sum f_i}{4} - F(\text{ant}) \right] \cdot h}{f_i}$	$Q_3 = l_i + \frac{\left[\frac{3 \cdot \sum f_i}{4} - F(\text{ant}) \right] \cdot h}{f_i}$
---	---	---

Exemplo:

i	Total de Pontos	f _i	F _i
1	150 - 154	4	4
2	154 - 158	9	13
3	158 - 162	11	24
4	162 - 166	8	32
5	166 - 170	5	37
6	170 - 174	3	40
	Total	40	

Primeiro Quartil

$$\frac{\sum f_i}{4} = \frac{40}{4} = 10, \text{ logo classe do } 1^\circ \text{ Quartil é } i = 2 \quad \ell = 154 \quad F(\text{ant}) = 4$$

h = 4 f₂ = 9

$$Q_1 = 154 + \frac{[10 - 4] \cdot 4}{9} = 154 + 2,66 = 156,66 = 156,7$$

Segundo Quartil = Mediana

$$\frac{2 \sum f_i}{4} = \frac{40}{2} = 20, \text{ logo classe do } 2^\circ \text{ Quartil é } i = 3 \quad \ell = 158 \quad F(\text{ant}) = 13$$

h = 4 f₃ = 11

$$Q_2 = Md = 158 + \frac{[20 - 13] \cdot 4}{11} = 158 + 2,5 = 160,5$$

Terceiro Quartil

$$\frac{3 \sum f_i}{4} = \frac{3 \cdot 40}{4} = 30, \text{ logo classe do } 3^\circ \text{ Quartil é } i = 4 \quad \ell = 162 \quad F(\text{ant}) = 24$$

h = 4 f₄ = 8

$$Q_3 = 162 + \frac{[30 - 24] \cdot 4}{8} = 162 + 3 = 165$$

Exercício: Calcule os quartis da seguinte distribuição:

i	Salário Mensal dos alunos do 3º Mat [R\$]	f _i	F _i
1	450 - 550	8	
2	550 - 650	10	
3	650 - 750	11	
4	750 - 850	16	
5	850 - 950	13	
6	950 - 1050	5	
7	1050 - 1150	1	
	Total	64	

5.6 Os Percentis

Denomina-se percentis os noventa e nove valores que separam uma série em 100 partes iguais. Indica-se da seguinte forma:

$P_1, P_2, P_3, \dots, P_{99}$

Note-se que: $P_{50} = Md$, $P_{25} = Q_1$ e $P_{75} = Q_3$

Calcula-se da mesma forma que os quartis, só que aplicando:

$\frac{k \sum f_i}{100}$, sendo k o número de ordem do percentil.

$$P_k = \ell_i + \frac{\left[\frac{k \cdot \sum f_i}{100} - F(\text{ant}) \right] \cdot h}{f_i}$$

Exemplo:

i	Total de Pontos	f_i	F_i
1	150 - 154	4	4
2	154 - 158	9	13
3	158 - 162	11	24
4	162 - 166	8	32
5	166 - 170	5	37
6	170 - 174	3	40
	Total	40	

Tem-se para o oitavo percentil:

$$k = 8 \Rightarrow \frac{8 \sum f_i}{100} = \frac{8 \cdot 40}{100} = 3,2, \text{ logo classe do } 8^\circ \text{ Percentil é } i = 1$$

$$\ell = 150$$

$$F(\text{ant}) = 0$$

$$h = 4$$

$$f_i = 4$$

$$P_8 = 150 + \frac{[3,2 - 0] \cdot 4}{4} = 150 + 3,2 = 153,2$$

Exercício: Calcule o percentil de ordem 20 da seguinte distribuição:

i	Salário Mensal dos alunos do 3º Mat [R\$]	f_i	F_i
1	450 - 550	8	
2	550 - 650	10	
3	650 - 750	11	
4	750 - 850	16	
5	850 - 950	13	
6	950 - 1050	5	
7	1050 - 1150	1	
	Total	64	

Capítulo 6 - Medidas de Dispersão ou de Variabilidade

6.1 Amplitude total (AT)

a) a amplitude total é a diferença entre o maior valor e o menor valor observado:

$$AT = x_{MÁX} - x_{MÍN}$$

Exemplo: 40, 45, 48, 52, 54, 62, e 70

$$AT = 70 - 40 = 30$$

Quanto maior a amplitude total, maior será a dispersão dos valores da variável em torno da média.

6.2 Variância (s^2) e Desvio Padrão (s)

São mais estáveis que a amplitude total, não sofrem tanto a interferência de valores extremos.

a) **para dados não agrupados**

A variância é a média aritmética dos quadrados dos desvios:

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{\sum f_i} = \frac{\sum (x_i - \bar{x})^2}{n}$$

A variância é um número em unidade quadrada em relação a média, por isso, definiu-se o desvio padrão como a raiz quadrada da variância.

O desvio padrão é a raiz quadrada da média aritmética dos quadrados dos desvios.

Para evitar o acúmulo de erro por arredondamento, simplifica-se o cálculo do desvio padrão com a seguinte:

$$\sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

que resulta em:

$$s = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

Obs: Quando calcula-se a variância ou o desvio padrão de uma população através de uma amostra dessa, deve-se substituir o denominador n por n-1.

Propriedades:

1ª: Somando-se (ou subtraindo-se) uma constante a (de) todos os valores de uma variável, o desvio padrão não se altera.

2ª.: Multiplicando-se todos os valores de uma variável por uma constante (diferente de zero), o desvio padrão fica multiplicado por essa constante.

Exemplo: Calcule o desvio padrão da seguinte série:

i	x_i	x_i^2
1	8	64
2	10	100
3	11	121
4	15	225
5	16	256
6	18	324
Total	78	1090

$$s = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} = \sqrt{\frac{1090}{6} - \left(\frac{78}{6}\right)^2} = \sqrt{181,67 - 169} = 3,56$$

b) para dados agrupados sem intervalos de classe: deve-se levar em conta as frequências.

$$s = \sqrt{\frac{\sum (f_i \cdot x_i^2)}{n} - \left(\frac{\sum (f_i \cdot x_i)}{n}\right)^2}$$

Exemplo:

i	Qtde de filhos que se deseja ter (x_i)	f_i	$f_i \cdot x_i$	$f_i \cdot x_i^2$
1	0	2	0	0
2	1	6	6	6
3	2	12	24	48
4	3	7	21	63
5	4	3	12	48
Total		30	63	165

$$s = \sqrt{\frac{\sum (f_i \cdot x_i^2)}{n} - \left(\frac{\sum (f_i \cdot x_i)}{n}\right)^2} = \sqrt{\frac{165}{30} - \left(\frac{63}{30}\right)^2} = \sqrt{5,5 - 4,41} = 1,04$$

Exercício: Determine o desvio padrão.

i	Qtde de cursos de extensão realizados por ano (x_i) pelos alunos do 3º Mat	f_i	$f_i \cdot x_i$	$f_i \cdot x_i^2$
1	1	2		
2	2	5		
3	3	8		
4	4	6		
5	5	3		
6	6	1		
Total		25		

c) **para dados agrupados com intervalos de classe:** também leva-se em conta as frequências e x_i é o ponto médio do intervalo de classe.

Exemplo:

i	Total de Pontos	x_i	f_i	$f_i x_i$	$f_i x_i^2$
1	150 154	152	4	608	92416
2	154 158	156	9	1404	219024
3	158 162	160	11	1760	281600
4	162 166	164	8	1312	215168
5	166 170	168	5	840	141120
6	170 174	172	3	516	88752
	Total		40	6440	1038080

$$s = \sqrt{\frac{\sum (f_i \cdot x_i^2)}{n} - \left(\frac{\sum (f_i \cdot x_i)}{n} \right)^2} = \sqrt{\frac{1038080}{40} - \left(\frac{6440}{40} \right)^2} = \sqrt{25952 - 25921} = \sqrt{31} = 5,57$$

Processo breve: Da mesma maneira que o cálculo da média, muda-se a variável X por outra Y, tal que:

$$y_i = \frac{x_i - x_0}{h} \quad \text{e} \quad s = h \cdot \sqrt{\frac{\sum (f_i \cdot y_i^2)}{n} - \left(\frac{\sum (f_i \cdot y_i)}{n} \right)^2}$$

Exemplo:

i	Total de Pontos	x_i	f_i	y_i	$f_i y_i$	$f_i y_i^2$
1	150 154	152	4	-2	-8	16
2	154 158	156	9	-1	-9	9
3	158 162	160	11	0	0	0
4	162 166	164	8	1	8	8
5	166 170	168	5	2	10	20
6	170 174	172	3	3	9	27
	Total		40		10	80

$$s = 4 \cdot \sqrt{\frac{80}{40} - \left(\frac{10}{40} \right)^2} = 4 \cdot \sqrt{2 - 0,0625} = 4 \cdot \sqrt{1,9375} = 4 \cdot 1,391941 = 5,57$$

Resolva: Calcule o desvio padrão pelo processo breve.

i	Salário Mensal dos alunos do 3º Mat [R\$]	x_i	f_i	y_i	$f_i y_i$	$f_i y_i^2$
1	450 550		8			
2	550 650		10			
3	650 750		11			
4	750 850		16			
5	850 950		13			
6	950 1050		5			
7	1050 1150		1			
	Total		64			

i	Peso kg	xi	fi	yi	f _i y _i	f _i y _i ²
1	30 - 50		2			
2	50 - 70		8			
3	70 - 90		12			
4	90 - 110		10			
5	110 - 130		5			
	Total		37			

6.3- Coeficiente de Variação (CV)

É a porcentagem do desvio padrão em relação a sua média.

$$CV = \frac{s}{\bar{x}} \cdot 100$$

Exemplo: Para o exemplo anterior, das estaturas, tem-se média de 161 cm e desvio padrão de 5,57 cm

$$CV = \frac{5,57}{161} \cdot 100 = 3,459 = 3,5\%$$

Resolva: Calcule o CV dos dois últimos exercícios de cálculo de desvio padrão pelo processo breve.

a) $\bar{x} = 755$
 $s = 154$

b) $\bar{x} = 84,3$
 $s = 21,88$

Conclusão: Quanto maior o CV maior será a dispersão
Quanto menor o CV menor será a dispersão

Exercícios de Revisão: Os dados abaixo referem-se a idade das pessoas que compraram um determinado produto novo durante um dia. Determine:

i	Idade	x_i	f_i	F_i	y_i	$f_i y_i$	$f_i y_i^2$	$f_i x_i$	$f_i x_i^2$
1	0 - 10		10						
2	10 - 20		26						
3	20 - 30		15						
4	30 - 40		8						
5	40 - 50		4						
6	50 - 60		3						
7	60 - 70		2						
	Total								

- a) Média;
- b) Desvio Padrão;
- c) Mediana
- d) Primeiro Quartil
- e) Terceiro Quartil
- f) P_{40}

TABELA DE NÚMEROS ALEATÓRIOS:

4 0 8 9 3 2 1 5 0 9 7 2 3 1 1 2 2 9 9 1 6 3 2 2 0 7 3 3 4 2 7 5 7 9 3 5
9 4 2 9 8 8 3 9 5 6 5 6 0 3 5 4 2 1 5 6 0 8 7 6 7 4 7 5 8 4 4 7 4 5 7 4
9 1 6 2 3 4 9 3 5 1 3 1 7 4 6 7 5 9 1 2 3 1 0 9 3 3 7 2 1 7 4 5 0 3 0 7
1 8 9 3 3 5 4 0 7 7 8 0 6 0 0 2 8 8 8 2 0 7 0 6 3 7 2 0 8 6 8 3 4 6 6 7
5 4 6 3 4 6 8 1 0 6 9 1 3 2 0 3 4 5 8 5 1 1 0 4 0 8 4 1 6 6 3 6 5 8 2 2
8 9 7 1 4 1 9 7 8 6 9 5 9 4 1 0 4 3 8 6 8 6 3 7 7 8 0 4 7 7 9 7 7 1 9 3
3 3 3 4 4 8 5 8 0 1 4 1 7 8 0 9 4 9 7 5 9 8 7 7 6 8 6 8 7 9 9 6 6 0 3 7
4 5 4 1 4 2 7 4 5 4 5 3 7 9 6 3 0 7 0 7 8 4 3 7 5 1 0 5 0 0 3 7 8 5 8 3
0 9 3 7 3 7 5 9 0 2 2 6 2 8 6 5 4 3 8 3 6 8 7 6 8 0 0 5 7 6 7 3 0 8 2 3
0 0 3 1 2 5 7 2 2 7 0 0 5 3 8 3 0 1 6 8 9 9 2 0 3 2 6 7 5 0 6 8 9 5 9 7
4 0 5 8 6 0 2 8 6 8 1 9 6 0 1 1 2 4 1 1 2 0 4 9 5 2 8 1 3 8 2 8 3 9 8 0
4 8 5 1 7 7 0 8 2 9 6 1 6 1 5 1 5 1 9 8 3 9 5 2 9 3 6 1 7 7 5 3 4 2 1 3
8 3 7 7 3 8 8 0 7 7 6 8 1 1 0 4 2 1 3 9 2 1 6 8 0 9 1 6 7 5 5 4 5 3 4 4
9 4 7 8 1 3 9 9 9 4 5 8 0 9 3 0 1 4 7 1 2 6 1 1 3 1 3 2 5 3 0 0 1 9 3 7
7 2 5 5 0 1 7 6 5 1 3 7 4 6 7 5 3 8 9 7 0 1 1 2 1 1 1 0 5 2 5 2 3 3 8 0
7 5 0 2 3 0 9 7 0 3 3 6 8 9 7 5 1 7 7 2 7 8 3 8 5 9 5 8 9 2 5 5 8 0 2 2
0 5 4 8 6 6 0 5 9 8 7 6 8 7 8 3 1 6 8 7 4 6 6 8 9 6 3 6 5 4 0 2 2 1 0 1
7 7 3 3 6 5 7 7 5 2 5 9 4 2 7 4 3 6 6 2 1 2 2 4 9 0 6 4 8 9 9 7 0 7 9 8
8 7 1 2 0 7 3 1 5 0 9 1 9 0 1 8 2 9 8 3 1 3 6 4 8 9 6 1 1 5 1 8 1 6 8 8
9 1 4 1 8 8 4 0 5 1 7 4 1 2 9 3 2 5 3 3 9 8 7 6 6 9 3 6 4 7 4 8 4 2 3 5
1 3 3 3 9 9 4 1 5 8 1 8 8 1 2 0 9 7 2 6 1 5 7 5 2 5 2 0 7 5 1 5 8 9 4 5
6 4 0 9 5 0 9 5 0 4 3 3 2 3 6 5 5 6 7 6 0 2 2 9 5 7 8 4 8 6 0 9 0 4 1 5
6 6 1 2 3 5 2 3 3 4 5 3 9 0 2 9 5 4 3 6 5 9 5 0 6 5 6 4 4 7 1 6 7 2 0 6
3 6 8 4 3 8 5 3 1 7 3 3 9 9 3 3 8 5 9 8 1 1 7 1 3 7 6 9 3 2 3 4 4 5 7 9
6 0 9 7 0 3 9 6 6 1 9 5 8 7 2 2 4 8 1 2 4 3 4 4 7 8 7 1 3 8 1 5 8 2 6 9
2 9 5 9 4 1 2 2 8 6 4 5 0 3 4 3 2 8 2 6 7 0 9 0 9 3 9 2 1 4 7 0 4 6 8 6
9 4 9 5 5 5 9 2 5 3 8 8 2 4 9 3 6 4 7 0 3 9 6 7 6 0 7 0 6 8 6 5 6 3 9 2
6 6 7 9 3 5 6 9 3 0 0 3 0 1 3 3 1 7 8 5 1 7 0 7 7 6 5 8 7 0 5 5 9 0 6 5
6 6 5 0 6 2 3 2 2 8 9 5 2 9 0 5 1 5 1 5 4 0 7 5 0 4 9 4 4 2 2 1 2 7 4 1
6 2 6 1 2 2 0 6 0 5 2 5 2 6 3 9 2 8 3 6 2 6 5 9 1 3 5 0 8 2 1 9 6 5 0 3
2 6 6 6 3 1 7 2 8 4 3 5 1 2 8 1 2 6 0 4 9 8 0 1 6 6 0 7 2 2 9 7 6 8 1 4
6 3 1 4 6 0 4 4 7 5 2 9 5 1 7 4 3 7 3 7 7 1 1 5 2 0 8 6 7 8 6 0 5 2 2 4
2 3 1 5 5 0 4 6 7 3 2 9 1 0 3 8 3 7 8 2 3 0 7 8 1 4 3 4 3 6 8 8 8 1 9 1
9 2 8 1 4 2 3 1 5 8 2 0 8 4 0 1 6 9 1 2 5 2 4 0 2 6 5 2 9 4 2 0 0 6 7 1
9 4 8 6 1 3 9 1 3 1 5 8 1 1 7 0 3 6 4 6 3 8 9 1 4 1 7 2 6 0 4 5 1 2 3 9
9 3 1 8 4 1 6 1 2 8 4 8 0 9 0 4 7 5 6 0 0 4 5 8 5 0 4 1 8 0 1 2 7 1 8 0
4 5 8 4 2 0 2 4 6 0 6 4 9 8 2 5 0 7 5 1 8 3 4 8 9 5 9 9 2 6 0 0 6 1 6 8
8 7 5 2 6 5 0 7 2 0 2 2 0 7 2 0 0 6 2 1 5 0 9 2 0 8 2 2 9 9 4 6 8 5 9 3
7 6 6 1 7 5 1 3 7 8 6 5 6 8 9 1 3 1 3 6 4 8 7 8 9 0 7 1 3 6 2 9 8 8 7 3
3 1 7 8 9 0 4 7 7 2 9 4 4 1 4 5 1 1 5 9 4 4 7 1 6 5 7 6 9 5 6 0 2 1 0 0
9 0 5 2 8 9 1 6 6 9 2 2 4 0 4 7 2 1 9 9 2 7 7 5 7 7 4 5 4 9 2 7 6 5 4 3
9 3 3 7 7 4 8 0 4 7 3 2 8 0 6 3 6 5 9 5 8 6 8 2 2 5 6 3 3 8 9 8 7 2 9 4
9 8 4 3 7 1 9 9 8 0 0 2 4 4 5 0 7 3 1 1 8 5 8 1 8 5 8 6 8 6 7 7 0 0 7 3
2 2 9 9 6 4 8 9 2 9 5 4 1 8 1 4 3 1 0 4 6 9 3 6 9 5 0 0 8 6 6 9 2 0 5 3
7 9 9 9 4 7 9 2 9 0 9 4 3 0 1 2 2 4 7 3 6 0 2 4 1 0 2 8 9 5 3 5 5 0 0 9
8 1 6 2 9 6 3 1 5 6 3 1 0 8 5 8 8 5 5 9 2 0 9 1 9 4 4 8 2 1 6 3 5 6 9 3
4 5 7 2 1 6 5 0 1 2 9 9 8 9 2 9 1 1 5 8 3 6 9 5 1 6 6 7 5 3 2 7 1 6 8 2
7 4 0 2 0 7 8 8 9 1 4 0 1 8 7 8 9 1 1 1 1 8 5 3 5 9 8 5 3 8 5 4 2 9 2 9
9 0 2 1 4 0 9 2 5 0 6 3 0 9 9 0 1 1 2 4 9 7 1 5 2 2 4 6 8 3 9 9 9 2 1 5
8 7 4 1 4 7 9 7 4 8 7 0 8 6 2 7 4 5 1 7 0 4 5 1 5 0 3 9 4 4 4 8 3 6 9 0
3 3 5 3 8 3 6 1 0 6 8 9 0 0 7 1 5 2 0 1 8 0 7 4 2 8 2 7 2 8 2 1 8 7 3 5
6 1 8 0 4 8 5 7 8 4 0 3 4 9 2 9 4 4 1 2 7 5 4 9 8 3 5 2 8 0 5 6 0 2 8 2
6 6 5 6 6 0 8 3 9 5 1 6 7 3 7 9 1 7 4 2 5 5 4 2 9 8 6 0 5 5 7 3 8 3 0 4
9 1 3 6 3 8 0 0 4 3 5 2 6 8 2 2 5 4 1 0 3 5 3 7 0 9 9 7 8 0 7 0 8 6 3 1
3 2 3 9 0 5 8 7 8 4 4 0 0 9 6 1 2 2 6 1 4 1 2 3 3 1 5 2 9 3 2 7 3 3 1 4
6 3 8 1 2 7 1 9 8 8 3 7 1 9 7 3 2 7 4 0 0 5 9 5 9 2 3 1 3 2 5 6 3 2 9 4

Tamanho da Amostra para populações finitas

$$n = \frac{z^2 \cdot (x/n) \cdot [1 - (x/n)] \cdot N}{(N-1) \cdot e^2 + z^2 \cdot (x/n) \cdot [1 - (x/n)]}$$

n = tamanho da amostra

N = tamanho da população

e = % de erro na forma unitária

z = intervalo de confiança, 1,96 para 95% de confiança (valor usual)

2,58 para 99% de confiança.

x/n = proporção esperada. O valor de n é máximo para x/n = 0,50

Resultando em:

$$n = \frac{1,96^2 \cdot 0,50 \cdot [1 - 0,50] \cdot N}{(N-1) \cdot e^2 + 1,96^2 \cdot 0,50 \cdot [1 - 0,50]} =$$

$$n = \frac{0,9604 \cdot N}{(N-1) \cdot e^2 + 0,9604}$$

Exemplo:

erro 2%	0,02
z=	1,96
x/n =	0,5

População	Amostra
100	96
200	185
300	267
400	343
500	414
600	480
700	542
800	600
900	655
1000	706
1100	755
1200	800
1300	844
1400	885
1500	923
1600	960
1700	996
1800	1029
1900	1061
2000	1091

População	Amostra
10000	1936
20000	2144
30000	2223
40000	2265
50000	2291
60000	2309
70000	2321
80000	2331
90000	2339
100000	2345

População	Amostra
100000	2345
200000	2373
300000	2382
400000	2387
500000	2390
600000	2391
700000	2393
800000	2394
900000	2395
1000000	2395

População	Amostra
1000000	2395
2000000	2398
3000000	2399
4000000	2400
5000000	2400
6000000	2400
7000000	2400
8000000	2400
9000000	2400
10000000	2400
115000000	2401

Cálculo do erro

$$e = z \cdot \sqrt{\frac{(x/n) \cdot [1 - (x/n)]}{n}} \quad \text{para população desconhecida}$$

$$e = z \cdot \sqrt{\frac{(x/n) \cdot [1 - (x/n)]}{n}} \cdot \sqrt{\frac{N - n}{N - 1}} \quad \text{para população conhecida}$$

para $z = 1,96$ e $x/n = 0,50$ tem-se:

$$e = 0,98 \cdot \sqrt{\frac{1}{n}} \quad \text{para população desconhecida}$$

$$e = 0,98 \cdot \sqrt{\frac{N - n}{n \cdot (N - 1)}} \quad \text{para população conhecida}$$

População =	100
Amostra	Erro
10	0,30
20	0,20
30	0,15
40	0,12
50	0,10
60	0,08
70	0,06
80	0,05
90	0,03
100	0,00

Bibliografia

STEVENSON, W. J. Estatística Aplicada à Administração. São Paulo: Editora HARBRA Ltda, 1981

BIBLIOGRAFIA:

COSTA NETO, P. L. de O. Probabilidades. São Paulo: Editora Edgard Blucher Ltda, 1985.

COSTA NETO, P. L. de O. Estatística. São Paulo: Editora Edgard Blucher Ltda, 17^o ed. 1999.

CRESPO, A. A. Estatística Fácil. São Paulo: Editora Saraiva, 17^o ed. 1999.

DANTE, L. R. Matemática: Contexto de Aplicações. São Paulo: Editora Ática, 1999.

DOWNING, D. , CLARK, J. Estatística Aplicada. São Paulo: Editora Saraiva, 2000.

KAZMIER, L. J. Estatística Aplicada à Economia e Administração. São Paulo: Editora Makron books Ltda., 1982.

LAPPONI, J. C. Estatística Usando Excel. São Paulo: Editora Lapponi, 2000.

LEVIN, J. Estatística Aplicada a Ciências Humanas, 2^a edição. São Paulo: Editora Harper & Row do Brasil Ltda, 1978.

NICK, E. , KELLNER, S. R. O. Fundamentos de Estatística para as Ciências do Comportamento. Rio de Janeiro: Editora Renes, 1971.

SIEGEL, S. Estatística Não Paramétrica. São Paulo: Editora McGraw-Hill do Brasil Ltda, 1975.

STEVENSON, W. J. Estatística Aplicada à Administração. São Paulo: Editora Harper & Row do Brasil Ltda, 1981.

TRIOLA, M. F. Introdução à Estatística. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 7^a ed. 1999.

Amostragem

Microsoft Excel - Estatística usando o excel.xls

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

Arial 10 = 77%

	A	B	C	D	E
1					
2	Função Aleatório = aleatorio()		Função Aleatórioentre = aleatorioentre(mínimo;maximo)		
3	gera um número aleatório entre 0 e 1		gera um número inteiro aleatório entre dois valores		
4	0,55		mínimo		
5	tecle F9 para gerar um novo número		máximo		
6			100		
7			200		
8	Função Aleatório = int(100*aleatorio())		tecle F9 para gerar um novo número		
9	gera um número aleatório entre 0 e 100				
10	69,00				
11	tecle F9 para gerar um novo número				
12					
13	Função Aleatório = int(1000*aleatorio())				
14	gera um número aleatório entre 0 e 1000				
15	201,00				
16	tecle F9 para gerar um novo número				
17					
18					
19					
20					
21					
22					
23					
24					
25					

Amostragem / Distribuição de Frequência / Dist.Binomial / Dis / NUM

Pronto

Distribuição de Frequência

Microsoft Excel - Estatística usando o excel.xls

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

Arial 10 = Seja a seguinte série de valores

	A	B	C	D	E	F
1	Seja a seguinte série de valores			Função máximo =máximo(início,fim)		
2	número de contratos vendidos			determina o maior valor de uma série		
3	14	13		17		
4	12	14				
5	13	15		Função mínimo =mínimo(início,fim)		
6	11	13		determina o menor valor de uma série		
7	12	12		11		
8	13	14				
9	16	13		Função cont.núm =cont.núm(início,fim)		
10	14	14		conta a quantidade de valores		
11	14	13		26		
12	15	15				
13	17	16				
14	14	12				
15	11	12				
16						
17						
18						

Amostragem Distribuição de Frequência Dist.Binomial Dis

Pronto NUM

Microsoft Excel - Estatística usando o excel.xls

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

Arial 10 =

115

1 Frequência Simples

2

3 Função frequência =frequência[série de dados;valor]

4 determina o número de vezes que o valor ocorre na série

5

6 1o passo: Pressionando o botão do mouse na célula J3 e arrastando até a J10, contendo uma linha mais]

7

8

9 2o passo: digite =frequência(A3:B15;3:9) e tecle control + shift + enter

10

11

12 Frequência Acumulada

13

14 digite =frequência[série;valor]

15

16 O valor zero final indica que nenhum valor da série foi excluído

17

18

19

20

21

22

23

24

xi	fi	Fi	fri	Fri
11	2	2	0,08	0,08
12	5	7	0,19	0,27
13	6	13	0,23	0,50
14	7	20	0,27	0,77
15	3	23	0,12	0,88
16	2	25	0,08	0,96
17	1	26	0,04	1,00
	0		1,00	

Amostragem Distribuição de Frequência Dist.Binomial Dis

Pronto NUM

Microsoft Excel - Estatística usando o excel.xls

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

Arial 10 =

	N	O	P	Q	R	S	T	U	V	W
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										

Pronto

Amostragem Distribuição de Frequência Dist.Binomial Dis