

Estatística

2 – Estatística Descritiva

Prof. Marcela A. G. Machado

Página da FEG: www.feg.unesp.br/~marcela

- Parâmetros de Posição:
 - Média
 - Mediana (2 partes)
 - Moda
- Parâmetros de Dispersão:
 - Variância
 - Desvio-Padrão
 - Coeficiente de Variação
 - Amplitude
- Parâmetros de Assimetria:
 - Coeficiente de Assimetria
- Parâmetros de Achatamento:
 - Coeficiente de Curtose

Média da População : μ

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

- Dados em tabela de freqüências:

$$\mu = \frac{\sum_{i=1}^k x_i \cdot f_i}{N} = \sum_{i=1}^k x_i \cdot p_i'$$

Onde:

x_i = valores da variável

f_i = freqüência da variável

N = tamanho da População

p_i' = freqüência relativa

k = número de linhas da tabela

- Dados em tabela com classes de freqüências:

x_i = pontos médios das classes.

Propriedades:

- Se $y_i = c \cdot x_i$ então $\mu_Y = c \cdot \mu_X$

- Se $y_i = x_i \pm c$ então $\mu_Y = \mu_X \pm c$

Média da Amostra ou Média Amostral: \bar{X}

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

- Dados em tabela de freqüências:

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n} = \sum_{k=1}^k x_i \cdot p'_i$$

Onde:

x_i : valores da amostra

f_i : freqüência

n : tamanho da amostra

$p'_i = f_i/n$: freqüência relativa

k : número de linhas da tabela

- Dados em tabela com classes de freqüências:

x_i : pontos médios das classes.

Média da Amostra ou Média Amostral: \bar{x}

Exemplo: *População:* Total de Funcionários de uma empresa

Amostra: 50 funcionários

Variável: Tempo gasto para preencher um formulário (s)

<i>Classes</i>				
Limites Aparentes		Limites Reais		
Primeira notação	Segunda notação		f_i	F_i
40 →45	40-44	39,5 - 44,5	3	3
45 →50	45-49	44,5 - 49,5	8	11
50 →55	50-54	49,5 - 54,5	16	27
55 →60	55-59	54,5 - 59,5	12	39
60 →65	60-64	59,5 - 64,5	7	46
65 →70	65-69	64,5 - 69,5	3	49
70 →75	70-74	69,5 - 74,5	1	50
			50	

Limites Reais		x_i	f_i	$x_i \cdot f_i$
39,5	44,5	42	3	126
44,5	49,5	47	8	376
49,5	54,5	52	16	832
54,5	59,5	57	12	684
59,5	64,5	62	7	434
64,5	69,5	67	3	201
69,5	74,5	72	1	72
		TOTAL	50	2725

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n} = \frac{2725}{50} = 54,5$$

Mediana (md)

Idéia: dividir o conjunto ordenado de valores em 2 partes

1 - Considerando conjunto com n valores ordenados:

• n (*ímpar*) \Rightarrow md = valor de ordem $(n + 1)/2$

Ex.:

35	36	37	38	40	40	41	43	46
----	----	----	----	----	----	----	----	----

n = 9 md = 40 (valor de ordem 5)

• n (*par*) \Rightarrow md = valor médio entre o de ordem $n/2$ e o de ordem $(n/2) + 1$

Ex.:

n = 8 valor de ordem $n/2 = 15$ valor de ordem $(n/2) + 1 = 16$

12	14	14	15	16	16	17	20
----	----	----	----	----	----	----	----

$$md = \frac{15 + 16}{2} = 15,5$$

Mediana (md)

2 - Considerando distribuição em classes de freqüências:

$$md = L_i + \frac{\left(\frac{n}{2} - F_a\right)}{f_{md}} \cdot h_{md}$$

onde:

L_i : limite inferior da classe que contém a mediana

n : número de elementos do conjunto de dados;

F_a : frequência acumulada das classes anteriores à classe que contém a mediana;

f_{md} : freqüência da classe que contém a mediana;

h_{md} : amplitude da classe que contém a mediana.

Exemplo: *População:* Total de Funcionários de uma empresa

Amostra: 50 funcionários

Variável: Tempo gasto para preencher um formulário (s)

Limites Reais		f_i	F_i	
39,5	44,5	3	3	
44,5	49,5	8	11	
49,5	54,5	16	27	
54,5	59,5	12	39	
59,5	64,5	7	46	
64,5	69,5	3	49	
69,5	74,5	1	50	
		50		

$$L_i = 49,5$$

$$n = 50$$

$$F_a = 11$$

$$f_{md} = 16$$

$$h_{md} = 5$$

$$md = 49,5 + \frac{\left(\frac{50}{2} - 11\right)}{16} \cdot 5 = 53,875$$

Moda (mo)

Valor de máxima freqüência dentro de um conjunto de dados

1 - Considerando conjunto com n valores

Exemplo: *População*: Peças produzidas numa empresa

Amostra: 20 peças

Variável: N defeitos por peça

X_i	f_i	p_i	
0	4	0,20	
1	7	0,35	← moda
2	5	0,25	
3	2	0,10	
4	1	0,05	
5	1	0,05	
	20	1,00	

2 - Considerando distribuição em classes de freqüências:

$$mo = L_i + \frac{d_1}{d_1 + d_2} \cdot h$$

Onde:

L_i : limite inferior da classe modal;

d_1 : diferença entre a freqüência da classe modal e a da classe imediatamente anterior;

d_2 : diferença entre a freqüência da classe modal e a da classe imediatamente seguinte;

h : amplitude da classe que contém a moda.

Moda (mo)

Exemplo: *População*: Total de funcionários de uma empresa

Amostra: 50 funcionários

Variável: Tempo gasto para preencher um formulário (s)

Limites Reais	fi	
39,5 - 44,5	3	
44,5 - 49,5	8	
49,5 - 54,5	16	← Classe modal
54,5 - 59,5	12	
59,5 - 64,5	7	
64,5 - 69,5	3	
69,5 - 74,5	1	
	50	

$$mo = L_i + \frac{d_1}{d_1 + d_2} \cdot h$$

$$L_i = 49,5$$

$$d_i = 16 - 8 = 8$$

$$d_2 = 16 - 12 = 4$$

$$h = 5$$

$$mo = 49,5 + \frac{8}{8 + 4} \cdot 5 = 52,8$$

Variância

Idéia: Média do quadrado da diferença dos dados em relação ao valor médio.

1 - Considerando dados de toda a População:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$$

2 - Considerando dados de uma Amostra da População:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Propriedades:

- **Se** $y_i = c \cdot x_i$ **Então:** $s_y^2 = c^2 \cdot s_x^2$
- **Se** $y_i = x_i \pm c$ **Então:** $s_y^2 = s_x^2$

Variância

Exemplo:

Seja os valores da amostra: 15 12 10 17 16

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

$$\bar{x} = \frac{70}{5} = 14$$

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
15	1	1
12	-2	4
10	-4	16
17	3	9
16	2	4
70	0	34

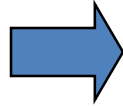
$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} = \frac{34}{5 - 1} = 8,5$$

Variância

Considerando distribuição de freqüências:

$$s_X^2 = \frac{\sum_{i=1}^k (x_i - \bar{x})^2 \cdot f_i}{n - 1}$$

equivalente



$$s_X^2 = \frac{\sum_{i=1}^k x_i^2 \cdot f_i - \left(\sum_{i=1}^k x_i \cdot f_i \right)^2 / n}{n - 1}$$

Exemplo: *População:* Total de Funcionários de uma empresa

Amostra: 50 funcionários

Variável: Tempo gasto para preencher um formulário (s)

Classes	f_i	x_i	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \cdot f_i$
39,5-44,5	3	42	156,25	468,75
44,5-49,5	8	47	56,25	450
49,5-54,5	16	52	6,25	100
54,5-59,5	12	57	6,25	75
59,5-64,5	7	62	56,25	393,75
64,5-69,5	3	67	156,25	468,75
69,5-74,5	1	72	306,25	306,25
	50	399	743,75	2262,5

$$s_X^2 = \frac{2262,5}{50 - 1} = 46,17$$

Desvio-Padrão

População:

$$\sigma_x = \sqrt{\sigma_x^2}$$

Amostra:

$$s_x = \sqrt{s_x^2}$$

Exemplo:

$$s_x^2 = 46,17$$

$$s_x = \sqrt{s_x^2} = \sqrt{46,17} = 6,79$$

Amplitude (R)

$$R = X_{\text{máx}} - X_{\text{min}}$$

Exemplo: *População:* Total de Funcionários de uma empresa

Amostra: 50 funcionários

Variável: Tempo gasto para preencher um formulário (s)

Classes	f_i	x_i
39,5-44,5	3	42
44,5-49,5	8	47
49,5-54,5	16	52
54,5-59,5	12	57
59,5-64,5	7	62
64,5-69,5	3	67
69,5-74,5	1	72
50		

$$x_{\text{max}} = 72 \quad \text{e} \quad x_{\text{min}} = 42 \quad \text{Assim:} \quad R = 72 - 42 = 30$$

Coeficiente de Variação (CV)

Idéia: relação entre o desvio-padrão amostral e a média amostral (%)

$$CV = \frac{S_x}{\bar{X}} \cdot 100$$

Regra empírica:

- $CV < 15\%$ \longrightarrow baixa dispersão
- $15\% < CV < 30\%$ \longrightarrow média dispersão
- $CV > 30\%$ \longrightarrow elevada dispersão

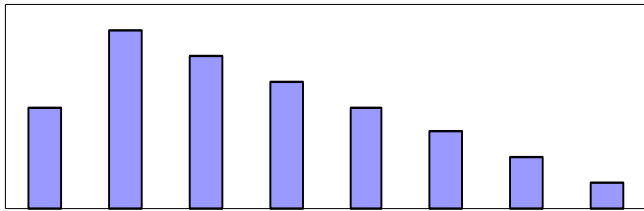
Exemplo anterior:

$$CV = \frac{S_x}{\bar{X}} \cdot 100 = \frac{6,79}{54,5} \cdot 100 = 12,5\%$$

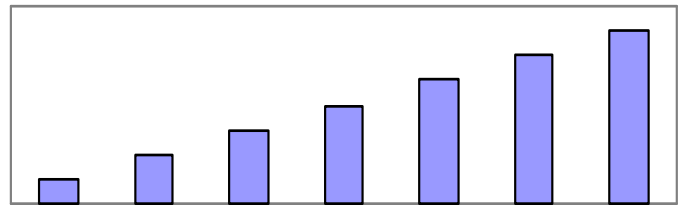
Medidas de Assimetria

Idéia: Como e quanto a distribuição de frequência se afasta da condição de simetria

Assimetria:



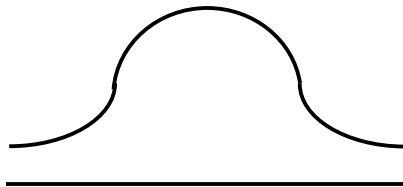
Assimetria Positiva



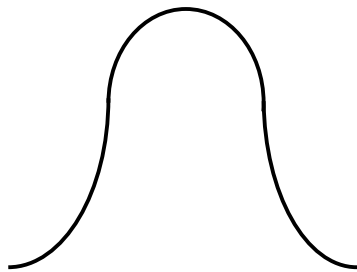
Assimetria Negativa

Medidas de Achatamento ou Curtose

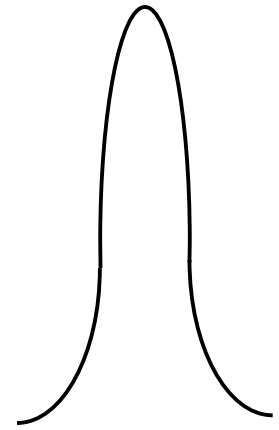
Idéia: Forma de distribuição quanto ao seu achatamento comparado com a distribuição Normal



Platicúrtica
(coef. curtose < 3)



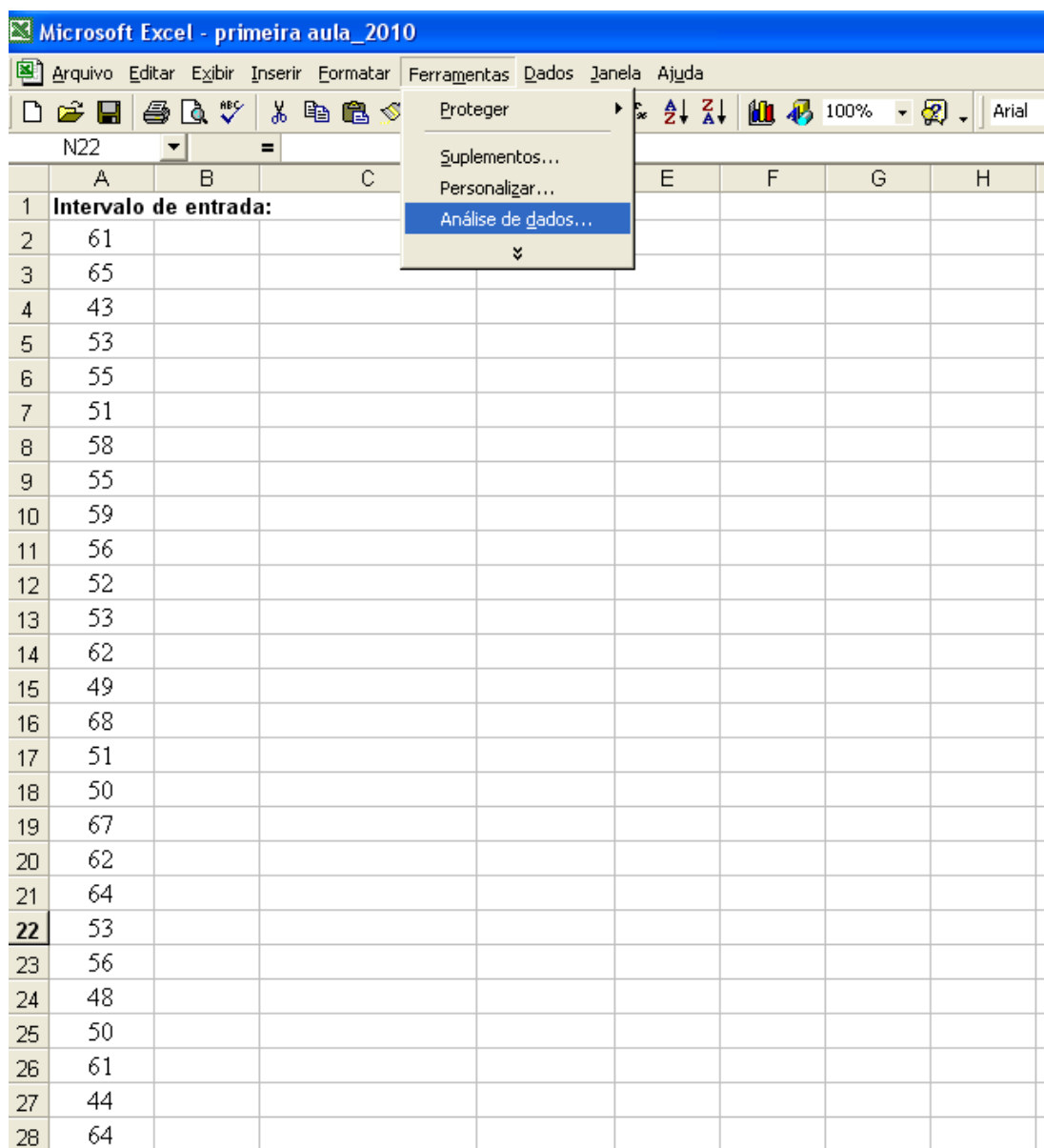
Mesocúrtica
(Normal)
(coef. curtose $= 3$)



Leptocúrtica
(coef. curtose > 3)

Estatística Descritiva - Excel

61	65	43	53	55	51	58	55	59	56
52	53	62	49	68	51	50	67	62	64
53	56	48	50	61	44	64	53	54	55
48	54	57	41	54	71	57	53	46	48
55	46	57	54	48	63	49	55	52	51



The screenshot shows the Microsoft Excel 2010 interface. The title bar reads 'Microsoft Excel - primeira aula_2010'. The menu bar includes 'Arquivo', 'Editar', 'Exibir', 'Inserir', 'Formatar', 'Ferramentas', 'Dados', 'Janela', and 'Ajuda'. The 'Ferramentas' menu is open, displaying options: 'Proteger', 'Suplementos...', 'Personalizar...', and 'Análise de dados...'. The 'Análise de dados...' option is highlighted. The spreadsheet area shows a column of data in column A, with row 1 labeled 'Intervalo de entrada:'. The data values in column A are: 61, 65, 43, 53, 55, 51, 58, 55, 59, 56, 52, 53, 49, 68, 51, 50, 67, 62, 64, 53, 56, 48, 50, 61, 44, 64, 53, 54, 55, 48, 54, 57, 41, 54, 71, 57, 53, 46, 48, 55, 46, 57, 54, 48, 63, 49, 55, 52, 51.

	A	B	C	E	F	G	H
1	Intervalo de entrada:						
2	61						
3	65						
4	43						
5	53						
6	55						
7	51						
8	58						
9	55						
10	59						
11	56						
12	52						
13	53						
14	62						
15	49						
16	68						
17	51						
18	50						
19	67						
20	62						
21	64						
22	53						
23	56						
24	48						
25	50						
26	61						
27	44						
28	64						

Estatística Descritiva - Excel

Microsoft Excel - primeira aula_2010

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

100% Arial

E13

	A	B	C	D	E	F	G	H	I	J
1	Intervalo de entrada:									
2	61									
3	65									
4	43									
5	53									
6	55									
7	51									
8	58									
9	55									
10	59									
11	56									
12	52									
13	53									
14	62									
15	49									
16	68									
17	51									
18	50									
19	67									
20	62									
21	64									
22	53									
23	56									
24	48									
25	50									
26	61									
27	44									
28	64									
29	53									
30	54									
31	55									
32	48									
33	54									
34	57									
35	41									

Análise de dados

Ferramentas de análise

- Anova: fator único
- Anova: fator duplo com repetição
- Anova: fator duplo sem repetição
- Correlação
- Covariância
- Estatística descritiva**
- Ajuste exponencial
- Teste-F: duas amostras para variâncias
- Análise de Fourier
- Histograma

OK Cancelar Ajuda

Estatística Descritiva - Excel

Microsoft Excel - primeira aula_2010

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

E13

	A	B	C	D	E	F	G	H	I	J
1	Intervalo de entrada:									
2	61									
3	65									
4	43									
5	53									
6	55									
7	51									
8	58									
9	55									
10	59									
11	56									
12	52									
13	53									
14	62									
15	49									
16	68									
17	51									
18	50									
19	67									
20	62									
21	64									
22	53									
23	56									
24	48									
25	50									
26	61									
27	44									
28	64									
29	53									
30	54									
31	55									
32	48									
33	54									
34	57									
35	41									
36	54									

Estatística descritiva

Entrada

Intervalo de entrada:

Agrupado por: ☒ Colunas ☐ Linhas

☐ Rótulos na primeira linha

Opções de saída

☒ Intervalo de saída:

☐ Nova planilha:

☐ Nova pasta de trabalho

☒ **Resumo estatístico**

☐ Nível de confiabilidade p/ média: %

☐ Enésimo maior:

☐ Enésimo menor:

OK Cancelar Ajuda

Estatística Descritiva - Excel

Microsoft Excel - primeira aula_2010									
Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda									
O29 =									
	A	B	C	D	E	F	G	H	I
1	Intervalo de entrada:								
2	61								
3	65								
4	43								
5	53		Coluna1						
6	55								
7	51		Média	54,62					
8	58		Erro padrão	0,93192362					
9	55		Mediana	54					
10	59		Modo	53					
11	56		Desvio padrão	6,589695109					
12	52		Variância da amostra	43,42408163					
13	53		Curtose	-0,04840093					
14	62		Assimetria	0,375126375					
15	49		Intervalo	30					
16	68		Mínimo	41					
17	51		Máximo	71					
18	50		Soma	2731					
19	67		Contagem	50					
20	62								
21	64								
22	53								
23	56								
24	48								
25	50								
26	61								
27	44								
28	64								
29	53								
30	54								
31	55								
32	48								
33	54								
34	57								
35	41								

Estatística Descritiva - Excel

Microsoft Excel - primeira aula_2010

ArquivoEditarExibirInserirFormatarFerramentasDadosJanelaAjuda

D6

1Intervalo de entrada

261

365

443

553

655

751

858

955

1059

1156

1252

1353

1462

1549

1668

1751

1850

1967

2062

2164

2253

2356

2448

2550

2661

2744

2864

2953

3054

3155

3248

3354

3457

3541

3654

Linhas

Colunas

Planilha

Gráfico...

Função...

</

Estatística Descritiva - Excel

Microsoft Excel - primeira aula_2010

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

D6

Intervalo de entrada:

	A	B	C	D	E	F	G	H	I
1	Intervalo de entrada:								
2	61								
3	65								
4	43								
5	53								
6	55			=					
7	51								
8	58								
9	55								
10	59								
11	56								
12	52								
13	53								
14	62								
15	49								
16	68								
17	51								
18	50								
19	67								
20	62								
21	64								
22	53								
23	56								
24	48								
25	50								
26	61								
27	44								
28	64								
29	53								
30	54								
31	55								
32	48								
33	54								
34	57								
35	41								

Colar função

Categoria da função: Nome da função:

Mais recentemente usada
Todas
Financeira
Data e hora
Matemática e trigonometria
Estatística
Procura e referência
Banco de dados
Texto
Lógica
Informações

INVLOG
INVT
LNGAMA
MAIOR
MÁXIMO
MÁXIMO A
MED
MÉDIA
MÉDIA.GEOMÉTRICA
MÉDIA.HARMÔNICA
MÉDIA.INTERNA

MÉDIA(núm1;núm2;...)

Retorna a média (aritmética) dos argumentos que podem ser números ou nomes, matrizes ou referências que contêm números.

OK Cancelar

Estatística Descritiva - Excel

Microsoft Excel - primeira aula_2010

Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda

MÉDIA = =MÉDIA(A2:A51)

	A	B	C	D	E	F	G	H	I
1	Intervalo de entrada:								
2	61								
3	65								
4	43								
5	53								
6	55			A(A2:A51)					
7	51								
8	58								
9	55								
10	59								
11	56								
12	52								
13	53								
14	62								
15	49								
16	68								
17	51								
18	50								
19	67								
20	62								
21	64								
22	53								
23	56								
24	48								
25	50								
26	61								
27	44								
28	64								
29	53								
30	54								
31	55								
32	48								
33	54								
34	57								
35	41								
36	54								

MÉDIA

Núm1 A2:A51 = {61;65;43;53;55;51

Núm2 = número

= 54,62

Retorna a média (aritmética) dos argumentos que podem ser números ou nomes, matrizes ou referências que contêm números.

Núm1: núm1;núm2;... de 1 a 30 argumentos numéricos cuja média se deseja obter.

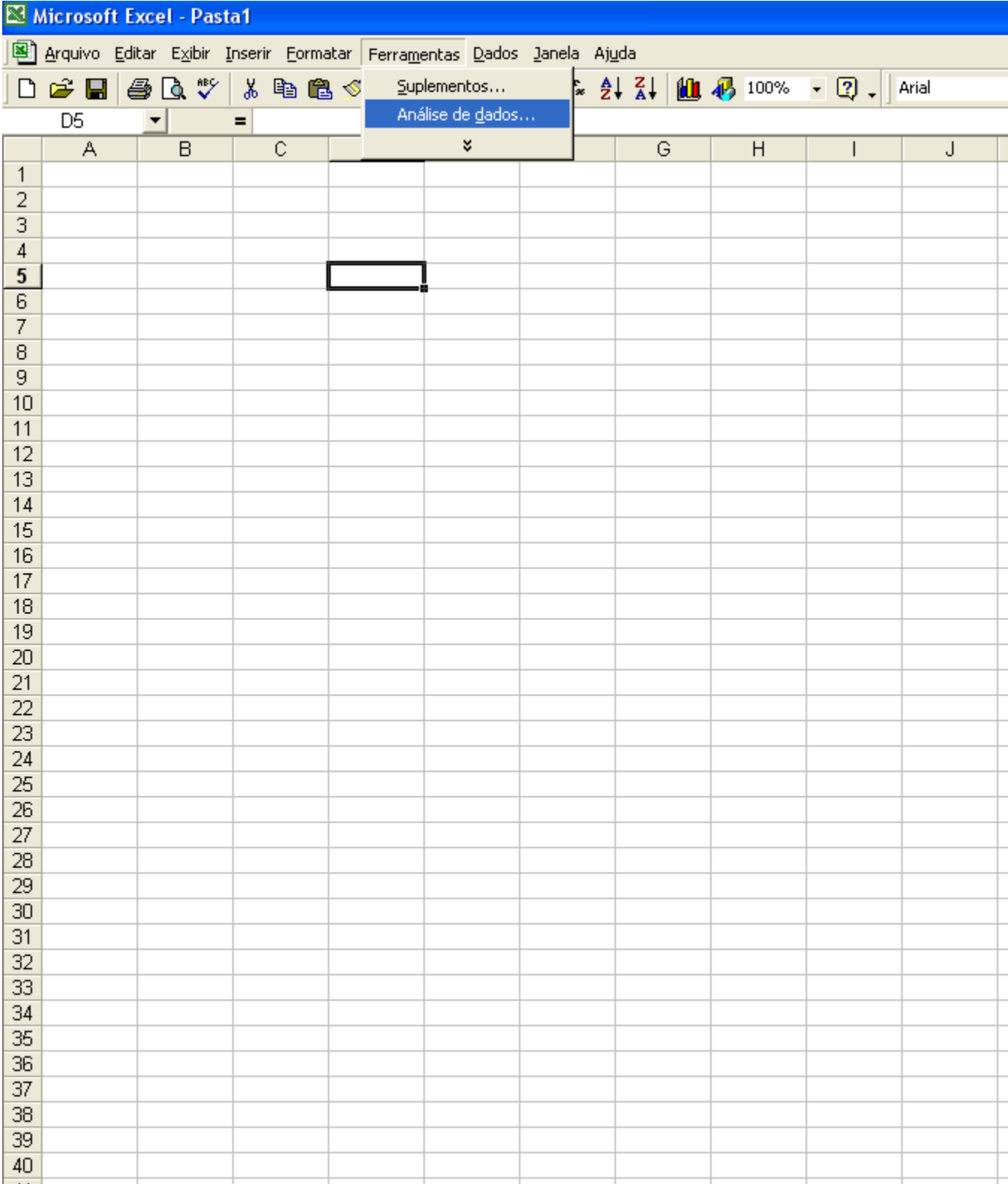
Resultado da fórmula = 54,62

OK Cancelar

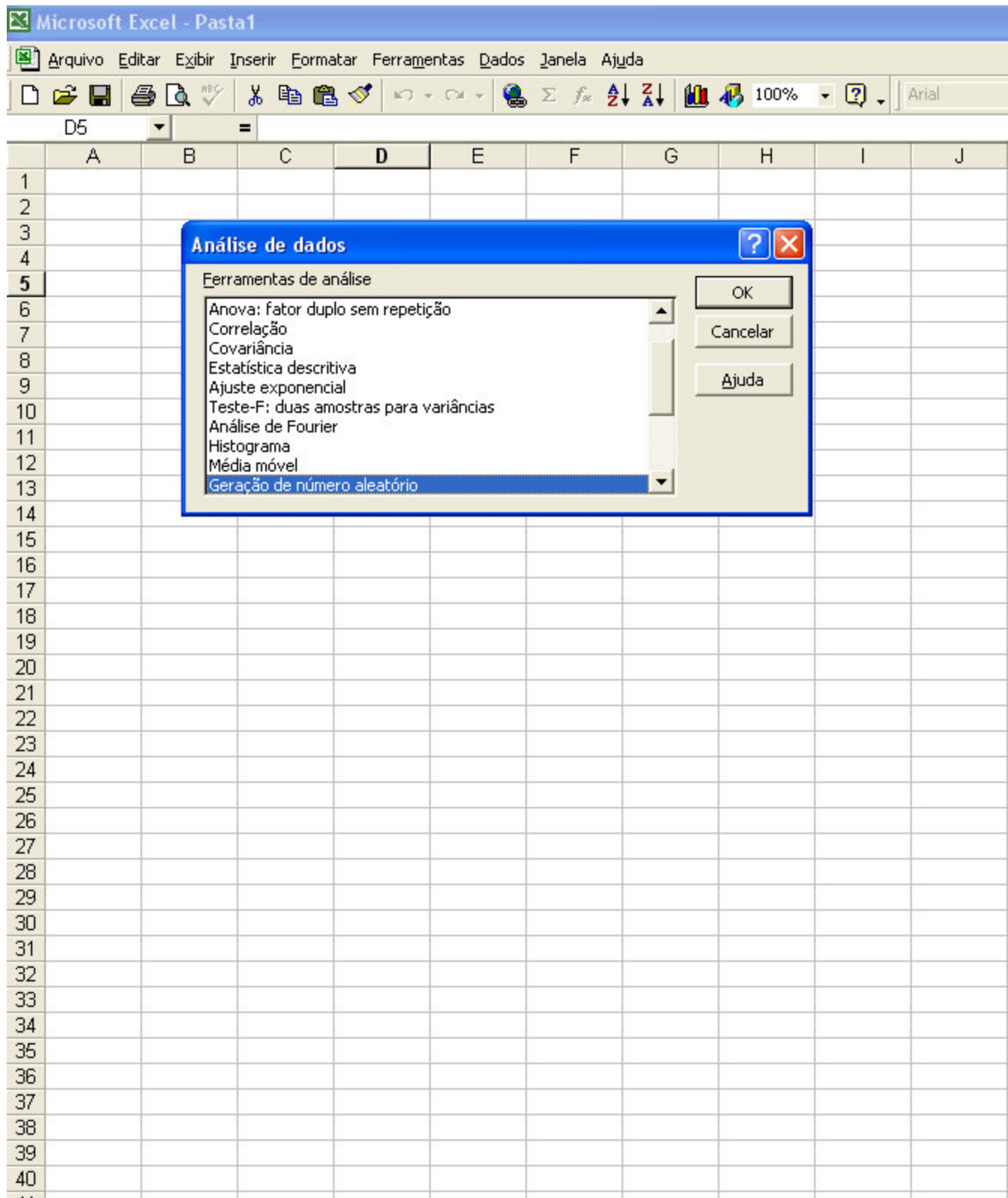
Estatística Descritiva - Excel

Microsoft Excel - primeira aula_2010									
Arquivo Editar Exibir Inserir Formatar Ferramentas Dados Janela Ajuda									
100% Arial									
D6	=MÉDIA(A2:A51)								
	A	B	C	D	E	F	G	H	I
1	Intervalo de entrada:								
2	61								
3	65								
4	43								
5	53								
6	55			54,62					
7	51								
8	58								
9	55								
10	59								
11	56								
12	52								
13	53								
14	62								
15	49								
16	68								
17	51								
18	50								
19	67								
20	62								
21	64								
22	53								
23	56								
24	48								
25	50								
26	61								
27	44								
28	64								
29	53								
30	54								
31	55								
32	48								
33	54								
34	57								
35	41								
36	54								

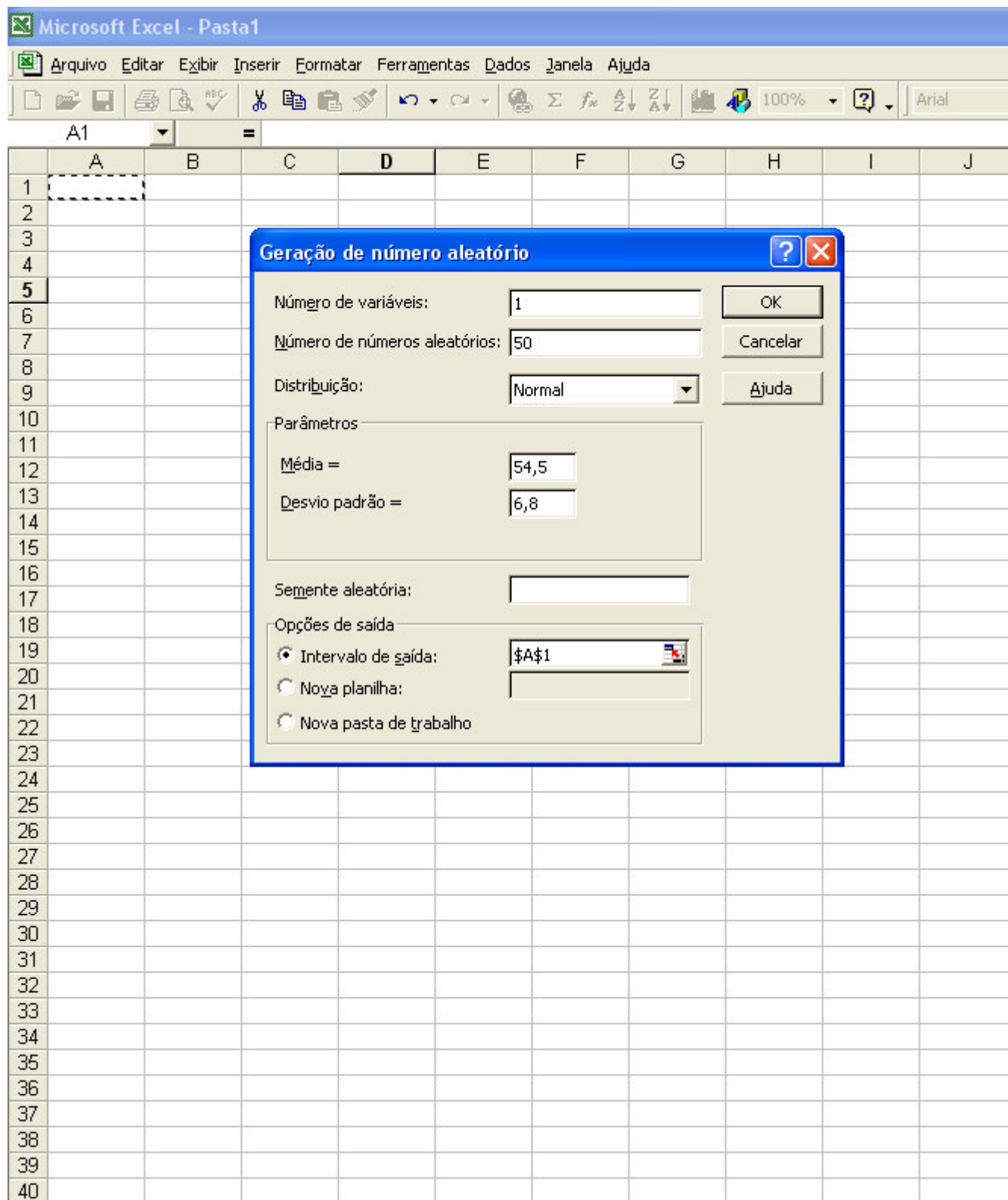
Geração de número aleatório - Excel



Geração de número aleatório - Excel



Geração de número aleatório - Excel



Geração de número aleatório - Excel

[illegible]