Estatística

2 - Estatística Descritiva

Possibilita descrever as Variáveis:

- ❑ DESCRIÇÃO GRÁFICA
- ❑ MEDIDAS DE POSIÇÃO
- ❑ MEDIDAS DE DISPERSÃO
- ❑ MEDIDAS DE ASSIMETRIA
- ❑ MEDIDAS DE ACHATAMENTO
- ❑ MEDIDAS DE CORRELAÇÃO

PROBLEMA:

Uma peça após fundida sob <u>pressão</u> a <u>alta temperatura</u> recebe um furo com diâmetro especificado em 12,00 mm e tolerância de 0,25 mm: (11,75 – 12,25)

Deseja-se DESCREVER as seguintes Variáveis de Resposta:

- **X: número de defeitos por peça fundida**
- **Y: diâmetro do furo**

Para tanto, coletou-se dados de uma Amostra de 25 peças

COLETA DE DADOS:

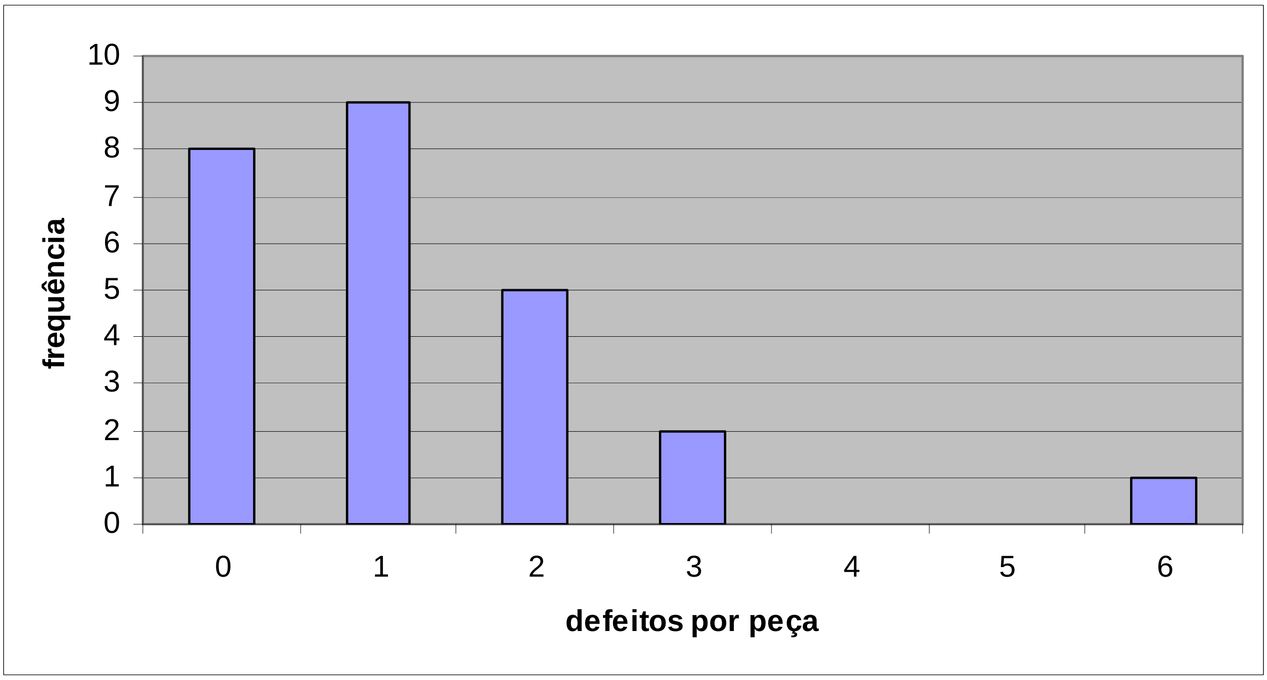
Peça i	X_i: número de defeitos	Y_i: diâmetro do Furo (mm)
1	2	12,21
2	0	11,73
3	1	11,94
4	2	11,86
5	1	12,31
6	0	12,10
7	1	12,19
8	0	11,78
9	2	12,20
10	1	12,05
11	1	11,81
12	3	12,00
13	1	12,34
14	0	11,99
15	2	12,27
16	1	12,11
17	**6**	11,80
18	3	12,02
19	0	12,23
20	1	12,08
21	0	11,88
22	1	11,76
23	2	12,05
24	0	12,07
25	0	12,20

VARIÁVEL X : Número de Defeitos por Peça

Tabela de Distribuição de frequências:

Ordem	X_i	f_i (absoluta)	p'_i (relativa)
		frequência	
1	0	8	32%
2	1	9	36%
3	2	5	20%
4	3	2	8%
5	4	0	0%
6	5	0	0%
7	6	1	4%
	total	25	100%

DIAGRAMA DE BARRAS (Variável Discreta)

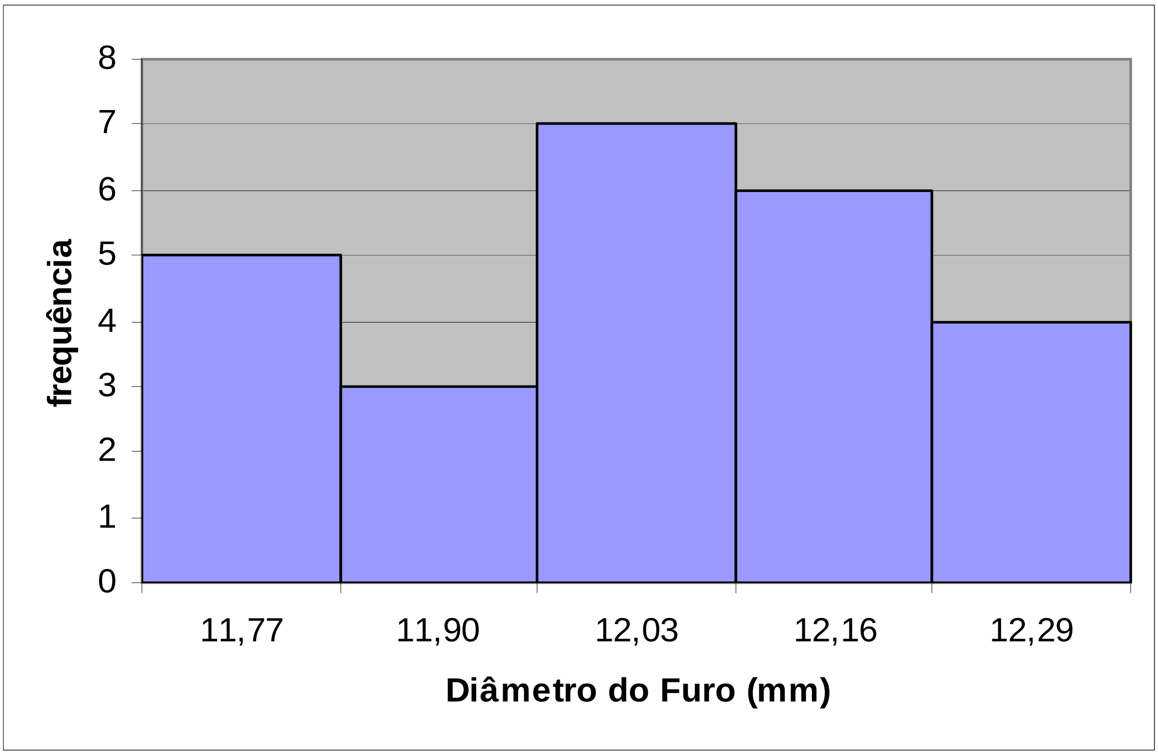


VARIÁVEL Y : Diâmetro de Furo (mm)

Tabela de Distribuição de frequências:

classe	Diâmetro do Furo	Valor médio Y_i	frequência f_i	frequência p_i'
1	11,705 até 11,835	11,77	5	20%
2	11,835 até 11,965	11,90	3	12%
3	11,965 até 12,095	12,03	7	28%
4	12,095 até 12,225	12,16	6	24%
5	12,225 até 12,355	12,29	4	16%
		total	25	100%

HISTOGRAMA (Variável Contínua)



HISTOGRAMA: dicas para construção

- Número de classes: $k = \sqrt{n}$ (inteiro)
- Amplitude da Amostra: $R = X_{max} - X_{min}$
- Amplitude das classes: $h = R/k$

- Exemplo da Fundição:

População: Total de peças produzidas

Tamanho da Amostra: n = 25 peças

Variável Y: diâmetro do furo (mm)

- Amplitude da amostra:

$$R = Y_{max} - Y_{min} = 12{,}34 - 11{,}73 = 0{,}61$$

- Número de classes:

$$k = \sqrt{n} = \sqrt{25} = 5$$

- Amplitude das classes:

$$h = \frac{R}{k} = \frac{0{,}61}{5} = 0{,}122 \rightarrow h = 0{,}13$$

HISTOGRAMA: dicas usando Excel:

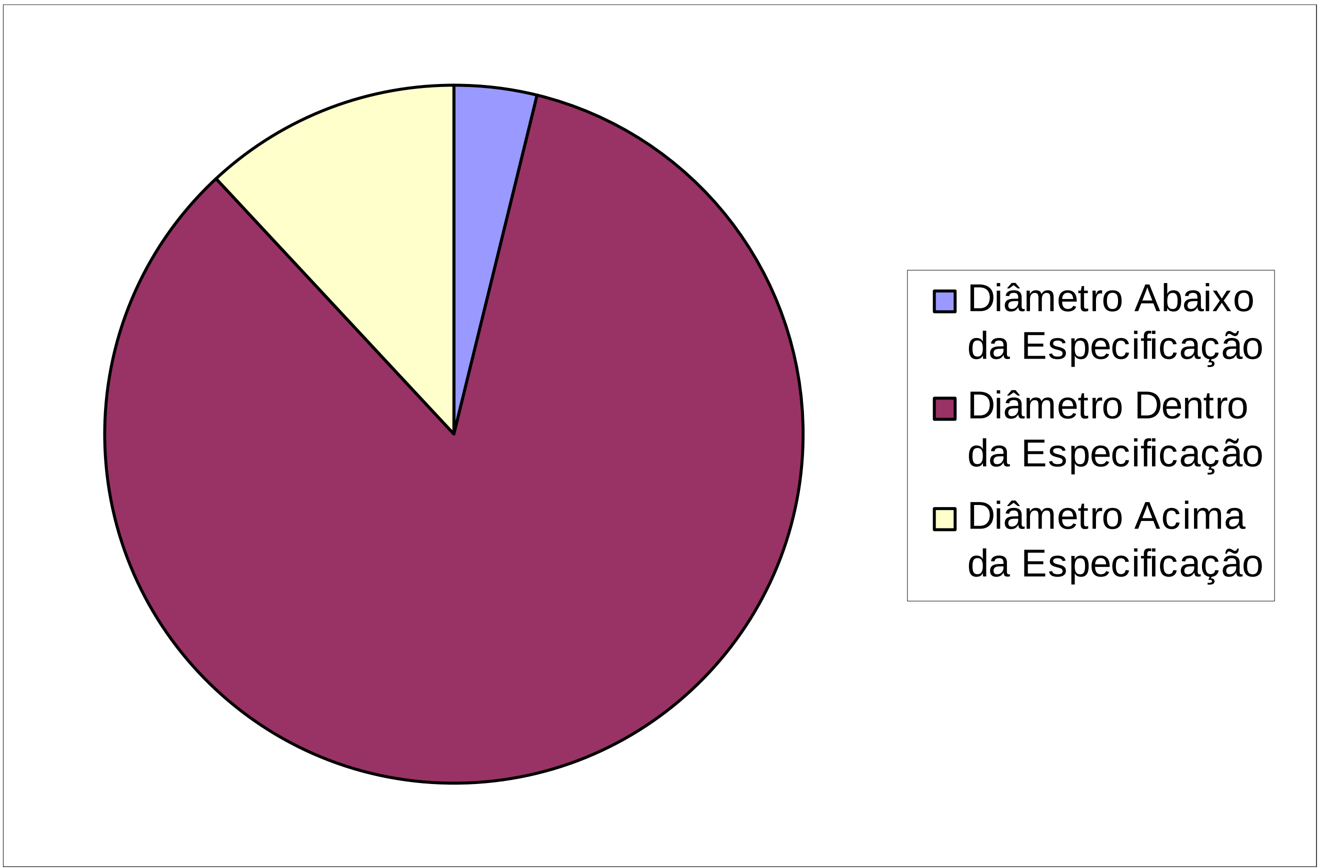
1. Selecionar: “Ferramentas”
   >> “Análise de Dados”
   >> “Histograma”
   >> “OK”
2. Selecionar células com os dados da Amostra
3. Selecionar células com os limites inferiores das classes
4. Escolher opção de saída
5. Selecionar “Porcentagem Cumulativa”
6. Selecionar “Resultado do Gráfico”
7. Clicar “OK”
8. Clicar em qualquer barra do Histograma gerado
9. Selecionar no painel superior: “Formatar”
   >> “Seqüência de dados selecionada”
   >> “Opções”
   >> “Espaçamento
   >> DIMINUIR PARA ZERO!
10. Clicar “OK”

VARIÁVEL: Categoria do Diâmetro de Furo

Distribuição de frequência:

classe	Diâmetro do Furo	Categoria	frequência absoluta	frequência relativa
1	< 11,75	abaixo da especificação	1	4%
2	11,75 até 12,25	dentro da especificação	21	84%
3	> 12,25	Acima da especificação	3	12%
		total	25	100%

DIAGRAMA CIRCULAR (PIZZA)



VARIÁVEL: Diâmetro de Furo (mm)

Peça i	Diâmetro Y_i
1	12,21
2	11,73
3	11,94
4	11,86
5	12,31
6	12,10
7	12,19
8	11,78
9	12,20
10	12,05
11	11,81
12	12,00
13	12,34
14	11,99
15	12,27
16	12,11
17	11,80
18	12,02
19	12,23
20	12,08
21	11,88
22	11,76
23	12,05
24	12,07
25	12,20

DIAGRAMA DE CAULE E FOLHAS

(Steam and Leaf Diagram)

11,7	3	6	8			
11,8	0	1	6	8		
11,9	4	9				
12.0	0	2	5	5	7	8
12,1	0	1	9			
12,2	0	0	1	3	7	
12,3	1	4				

MEDIDAS DE POSIÇÃO

- Média
- Mediana
- Quartil
- Decil
- Percentil
- Moda

Média da População (Variável X): E(X)

$$\boxed{\mu_X = \frac{\sum_{i=1}^{N} X_i}{N}}$$

X_i : i-ésimo valor da Variável X

N : tamanho da População

$\boxed{\mu_X}$ é um PARÂMETRO,
isto é, um DETERMINADO NÚMERO,
pois considera TODOS os possíveis
valores da População

Média da Amostra ou Média Amostral : $\overline{X}$

$$\overline{X} = \frac{\sum_{i=1}^{N} x_i}{n}$$

X_i : i-ésimo valor de uma Amostra da Variável X

n : tamanho da Amostra

$\overline{X}$ é uma VARIÁVEL,
pois depende dos valores de
cada Amostra

Dica “Excell” para a Média:
Selecionar: “fx” >> “Estatística” >> “Média” Selecionar: células com a tabela de dados Clicar: “OK”

Média da Amostra ou Média Amostral: $\overline{X}$

• Dados em Tabela de frequência dos valores de uma dada Amostra da Variável X

$$\boxed{\overline{X} = \frac{\sum_{i=1}^{k} x_i \cdot f_i}{n} = \sum_{i=1}^{k} x_i \cdot p_i^{'}}$$

f_i : frequência do valor X_i

$n = \sum_{i=1}^{k} f_i$: tamanho da Amostra

$p_i^{'} = f_i / n$: frequência relativa

k : número de diferentes valores da Amostra

• Dados em Tabela de frequência das classes de uma dada Amostra da Variável X

$$\boxed{\overline{X} = \frac{\sum_{i=1}^{k} x_i \cdot f_i}{n}}$$

x_i : valor médio da classe i

f_i : freqüência da classe i

k : número de classes

MÉDIA AMOSTRAL: Exemplo da Fundição

- Variável X: número de defeitos por peça

Tabela de Distribuição de frequência dos Valores

Ordem i	Número de Defeitos X_i	frequência f_i	$X_i \cdot f_i$
1	0	8	0
2	1	9	9
3	2	5	10
4	3	2	6
5	4	0	0
6	5	0	0
7	6	1	6
	total	25	31

$$\boxed{\overline{X} = \frac{\sum_{i=1}^{k} X_i \cdot f_i}{n} = \frac{31}{25} = 1{,}24}$$

MÉDIA AMOSTRAL: Exemplo da Fundição

- Variável Y: diâmetro do furo (mm)

Tabela de Distribuição de frequência das Classes

Classe i	Diâmetro do Furo	Y_i	frequência f_i	$Y_i \cdot f_i$
1	11,705 até 11,835	11,77	5	58,85
2	11,835 até 11,965	11,90	3	35,7
3	11,965 até 12,095	12,03	7	84,21
4	12,095 até 12,225	12,16	6	72,96
5	12,225 até 12,355	12,29	4	49,16
		total	25	300,88

$$\overline{Y} = \frac{\sum_{i=1}^{k} Y_i \cdot f_i}{n} = \frac{300{,}88}{25} = 12{,}04$$

Dica "Excell" para Média em tabela de frequência das classes:

Selecionar: "fx" >> "Matemática" >> **"SOMARPRODUTOS"**
Selecionar: células com os valores de X_i
Selecionar: células com os valores de f_i >>
Clicar: "OK" >> Dividir por n

MEDIANA : md

Idéia: dividir em 2 partes um conjunto ordenado de valores

1 - Tabela com n valores ordenados:

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

Exemplo:

ordem	1	2	3	4	5	6	7	8	9
valor	35	36	37	38	40	40	41	43	46

n = 9 (n+1)/2 = 5

valor de ordem 5 = 40

$$\boxed{md = 40}$$

❑ n: *par* $\Rightarrow$ md = valor médio entre o de ordem n/2 e o de ordem n/2+1

Exemplo:

ordem	1	2	3	4	5	6	7	8
valor	12	14	14	15	16	16	17	20

n = 8

valor de ordem n/2 = 15

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

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

Dica "Excell":
Selecionar: "fx" >> "Estatística" >> "Med"
Selecionar: células com a tabela de dados >> Clicar: "OK"

MEDIANA : md

2 – Tabela de Distribuição em classes de freqüências:

$$md = L_{md} + \frac{\left(\frac{n}{2} - F_{<md}\right)}{f_{md}} \cdot h$$

onde:

L_{md} : limite inferior da classe que contém a mediana

n : tamanho da Amostra

$F_{<md}$: frequência acumulada das classes anteriores à classe que contém a mediana

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

h : amplitude das classes

MEDIANA : md

Exemplo da Fundição:

Variável Y: diâmetro do furo (mm)

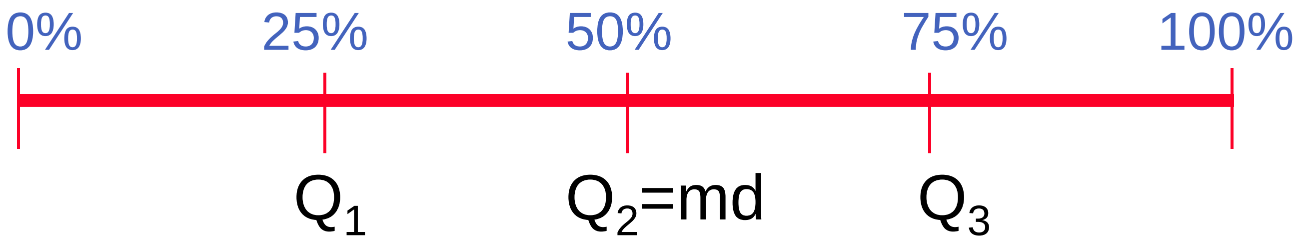
classe	Limites Reais		frequência	
	Lim. inf.	Lim. sup.	absoluta f_i	Acumulada F_i
1	11,705	11,835	5	5
2	11,835	11,965	3	**8**
3	**11,965**	12,095	**7**	15
4	12,095	12,225	6	21
5	12,225	12,355	4	**25**

$$\boxed{md = L_{md} + \frac{\left(\frac{n}{2} - F_{<md}\right)}{f_{md}} \cdot h}$$

$$\boxed{md = 11{,}965 + \frac{\left(\frac{25}{2} - 8\right)}{7} \cdot 1{,}30 = 12{,}80}$$

QUARTIL : Q_i

Idéia: dividir em 4 <u>partes</u> um conjunto ordenado de valores numéricos



Q_1: Primeiro Quartil

Q_2: Segundo Quartil = Mediana

Q_3: Terceiro Quartil

<u>Dica “Excell” (os dados não precisam estar ordenados):</u>
Selecionar:Inserir>>função “fx” >> “Estatística” >>“Quartil” >> OK
Selecionar: células com a tabela de dados >> OK
Seguir instruções da janela

QUARTIL : Q_i

Exemplo da Fundição:

Variável X: número de defeitos por peça

ordem i	X_i
1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	2
19	2
20	2
21	2
22	2
23	3
24	3
25	6

→ $Q_1 = 0$ (primeiro quartil)

→ $Q_2 = 1$ (segundo quartil)

→ $Q_3 = 2$ (terceiro quartil)

QUARTIL : Q_i

Exemplo da Fundição:

Variável Y: diâmetro do furo (mm)

peça i	Y_i
1	11,73
2	11,76
3	11,78
4	11,80
5	11,81
6	11,86
7	11,88
8	11,94
9	11,99
10	12,00
11	12,02
12	12,05
13	12,05
14	12,07
15	12,08
16	12,10
17	12,11
18	12,19
19	12,20
20	12,20
21	12,21
22	12,23
23	12,27
24	12,31
25	12,34

$\rightarrow$ Q_1= 11,88 (primeiro quartil)

$\rightarrow$ Q_2= 12,05 (segundo quartil)

$\rightarrow$ Q_3= 12,20 (terceiro quartil)

QUARTIL : Q_i

Distribuição em classes de freqüências:

$$Q_i = L_{Qi} + \frac{\left(\frac{i \cdot n}{4} - F_{<Qi}\right)}{f_{Qi}} \cdot h$$

onde:

L_{Qi} : limite inferior da classe que contém o i-ésimo Quartil

n: tamanho da Amostra

$F_{<Qi}$: frequência acumulada das classes anteriores à classe que contém o i-ésimo Quartil;

f_{Qi} : freqüência da classe que contém o i-ésimo Quartil;

H_i: amplitude das classes

QUARTIL : Q_i

Exemplo da Fundição:

Variável Y: diâmetro do furo (mm)

classe	Limites Reais		frequência	
	Lim. inf.	Lim. sup.	absoluta f_i	Acumulada F_i
1	11,705	11,835	5	5
2	11,835	11,965	3	8
3	11,965	12,095	7	15
4	12,095	12,225	6	21
5	12,225	12,355	4	25

Q_1 = valor de ordem 7 (25/4) $\Rightarrow$ classe 2

Q_2 = valor de ordem 13 (50/4) $\Rightarrow$ classe 3

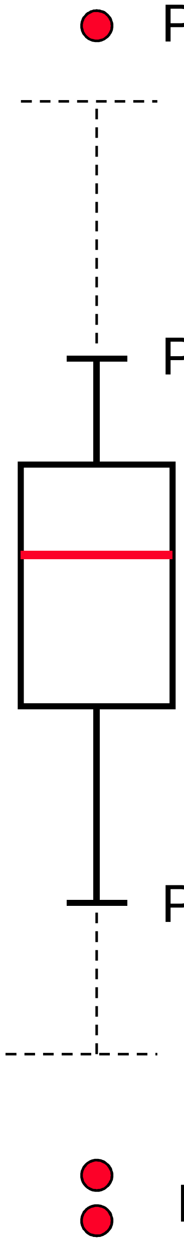
Q_3 = valor de ordem 19 (75/4) $\Rightarrow$ classe 4

$$\boxed{Q_i = L_{Qi} + \frac{\left(\frac{i \cdot n}{4} - F_{<Qi}\right)}{f_{Qi}} \cdot h}$$

$$\boxed{Q1 = 11{,}835 + \frac{\left(\frac{1 \cdot 25}{4} - 5\right)}{3} \cdot 0{,}13 = 11{,}89}$$

Analogamente: Q_2=12,05 Q_3=12,18

DIAGRAMA DE CAIXA E BIGODE *(Box-plot)*



PE: Ponto Extremo *(outlier)*

BS: Barreira Superior = $Q_3 + 1{,}5 \cdot (Q_3 - Q_1)$

PS: Ponto Adjacente Superior

Q_3: Terceiro Quartil

Q_2: Segundo Quartil = Mediana

Q_1: Primeiro Quartil

PS: Ponto Adjacente Inferior

BI: Barreira Inferior = $Q_1 - 1{,}5 \cdot (Q_3 - Q_1)$

PE: Pontos Extremos *(outliers)*

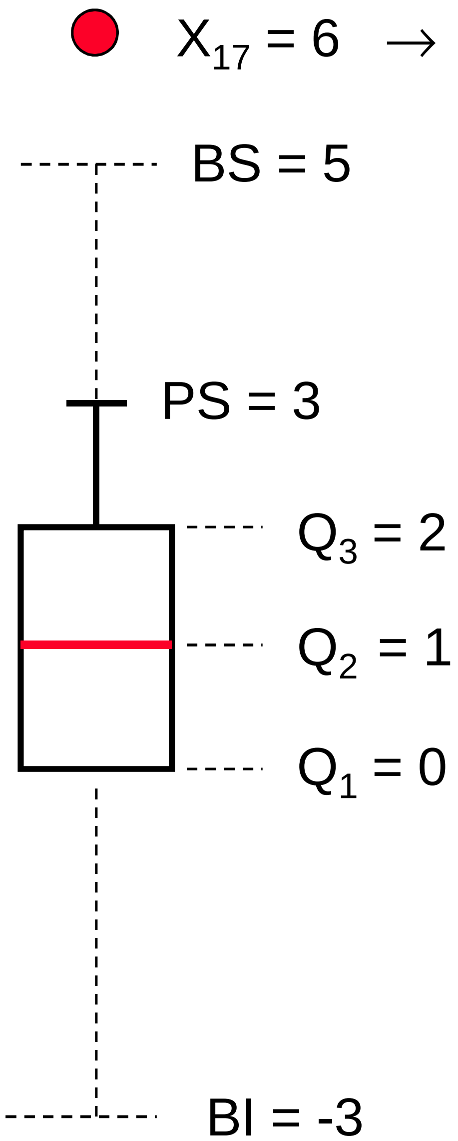
Usualmente:
- ✓ apresentar os Pontos Extremos
- ✓ não apresentar as Barreiras ((BI e BS)

DIAGRAMA DE CAIXA E BIGODE *(Box-plot)*

Exemplo da Fundição:

Variável X: número de defeitos por peça

$X_{17} = 6$ → Ponto Extremo (*outlier*)



$$BS = Q_3 + 1{,}5 \cdot (Q_3 - Q_1)$$

$$BS = 2 + 1{,}5 \cdot (2 - 0) = 5$$

PS = 0 (sem bigode inferior)

$$BI = Q_1 - 1{,}5 \cdot (Q_3 - Q_1)$$

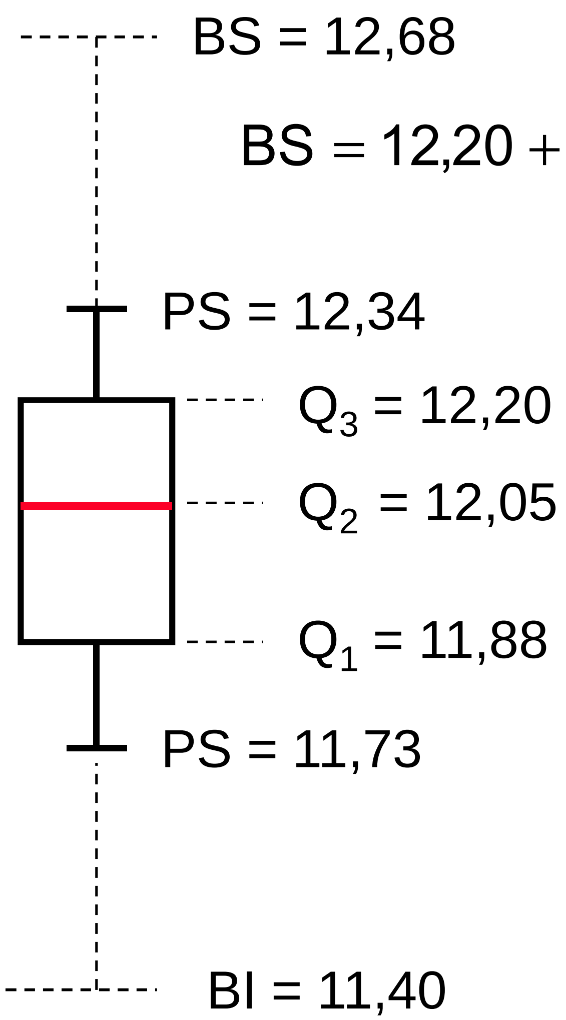
$$BI = 0 - 1{,}5 \cdot (2 - 0) = -3$$

DIAGRAMA DE CAIXA E BIGODE *(Box-plot)*

Exemplo da Fundição: Variável Y: diâmetro do furo (mm)

$$BS = Q_3 + 1{,}5 \cdot (Q_3 - Q_1)$$

$$BS = 12{,}20 + 1{,}5 \cdot (12{,}20 - 11{,}88) = 12{,}68$$



$$BI = Q_1 - 1{,}5 \cdot (Q_3 - Q_1)$$

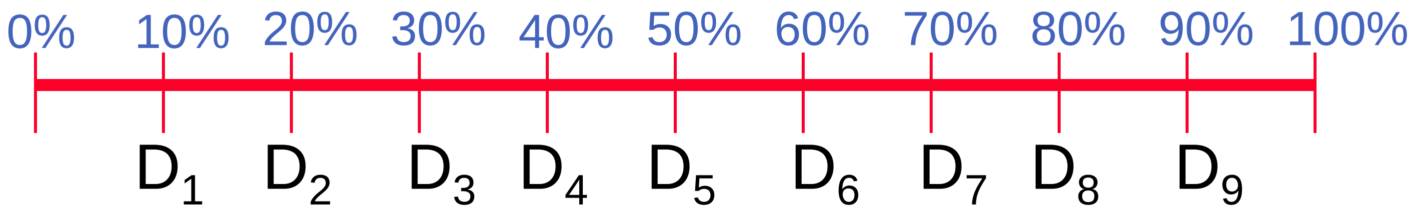
$$BI = 11{,}88 - 1{,}5 \cdot (12{,}20 - 11{,}88 = 11{,}40)$$

Observação:

✓ no exemplo não ocorreram Pontos Extremos

Decil (D_i)

Idéia: Dividir o conjunto de dados em 10 partes iguais



D_5= mediana

$$D_i = L_{Di} + \frac{\left(\frac{i \cdot n}{10} - F_{Di}\right)}{f_{Di}} \cdot h_{Di}$$

onde:

L_{Di}: limite inferior da classe que contém o i-ésimo Decil

n: número de elementos do conjunto de dados;

F_{Di}: frequência acumulada das classes anteriores à classe que contém o i-ésimo Decil;

f_{Di}: freqüência da classe que contém o i-ésimo Decil;

h_{Di}: amplitude da classe que contém o i-ésimo Decil.

Dica Excell: Como $D_i = P_{10.i}$ então pode-se usar a função “Percentil”

Percentil (P_i)

Idéia: Dividir o conjunto de dados em 100 partes iguais

0% 1% 2% 3% ... 50% ... 97% 98% 99% 100%

P_1 P_2 P_3 $P_{50}=md$ P_{97} P_{98} P_{99}

$$\boxed{P_i = L_{Pi} + \frac{\left(\frac{i \cdot n}{100} - F_{Pi}\right)}{f_{Pi}} \cdot h_{Pi}}$$

onde:

L_{Pi}: limite inferior da classe que contém o i-ésimo Percentil

n: número de elementos do conjunto de dados;

F_{Pi}: frequência acumulada das classes anteriores à classe que contém o i-ésimo Percentil

f_{Pi}: freqüência da classe que contém o i-ésimo Percentil

h_{Pi}: amplitude da classe que contém o i-ésimo Percentil

Dica “Excell”:

Selecionar: “fx” >> “Estatística” >> “Percentil”

Selecionar: células com a tabela de dados

Clicar: “OK” >> seguir instruções da janela

Moda: mo

Valor de máxima freqüência dentro de um conjunto de dados

- **Dados em Tabela de frequência dos valores**

  Exemplo da Fundição

  Variável X: número de defeitos por peça

Ordem	x_i	frequência f_i (absoluta)	frequência p_i' (relativa)
1	0	8	32%
2	**1**	**9**	**36%**
3	2	5	20%
4	3	2	8%
5	4	0	0%
6	5	0	0%
7	6	1	4%
	total	25	100%

mo = 1

moda é apresentar 1 defeito por peça

Moda: mo

- ❑ Dados em Tabelas de frequência das classes

Classe Modal: aquela(s) de maior frequência

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

Li : limite inferior da classe modal

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

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

h : amplitude das classes

Dica "Excell" (os dados não precisam estar ordenados):
Selecionar:Inserir>>função "fx" >> "Estatística" >>"Modo" >> OK
Selecionar: células com a tabela de dados >> OK
Seguir instruções da janela

Moda: mo

- **Dados em Tabelas de frequência das classes**

Exemplo da Fundição

Variável Y: diâmetro do furo (mm)

classe	Limites Reais		frequência	
	Lim. inf.	Lim. sup.	absoluta f_i	Acumulada F_i
1	11,705	11,835	5	5
2	11,835	11,965	3	8
3	11,965	12,095	7	15
4	12,095	12,225	6	21
5	12,225	12,355	4	25

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

$$\boxed{d_1 = 7 - 3 = 4}$$

$$\boxed{d_2 = 7 - 6 = 1}$$

$$\boxed{mo = 11{,}965 + \frac{4}{4+1} \cdot 0{,}13 = 12{,}07}$$

- Variância
- Desvio padrão
- Coeficiente de Variação
- Amplitude

Variância da População (Variável X): Var(X)

$$\boxed{\sigma_X^2 = \frac{\sum_{i=1}^{N}\left[X_i - \mu_x\right]^2}{N}}$$

X_i : i-ésimo valor da Variável X

μ_x : Média da População

N : tamanho da População

σ_X^2 é um PARÂMETRO,

isto é, um DETERMINADO NÚMERO,

pois considera TODOS os possíveis

valores da População

Variância da Amostra ou Variância Amostral : s_X^2

$$\boxed{S_X^2 = \frac{\sum_{i=1}^{N}\left[X_i - \overline{X}\right]^2}{n-1}}$$

X_i : i-ésimo valor de uma Amostra da Variável X

n : tamanho da Amostra

No Cap. 9 – Estimação de Parâmetros por Ponto
será apresentada a justificativa da divisão por **n-1**

S_X^2 é uma VARIÁVEL,
pois depende dos valores de
cada Amostra

Dica “Excell”:
Selecionar: “fx” >> “Estatística” >> “VAR”
Selecionar: células com a tabela de dados
Clicar: “OK”

Obs.: No caso de População usar “VARP”

Variância Amostral

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

Exemplo:

valores da amostra: 15 12 10 17 16

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

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

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

Variância Amostral

Considerando distribuição de freqüências de valores

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

equivalente ⇒

$$S_X^2 = \frac{\sum_{i=1}^{n} x_i^2 \cdot f_i - (\sum_{i=1}^{n} x_i \cdot f_i)^2 / n}{n - 1}$$

Exemplo da Fundição X: número de defeitos por peça

	X_i	f_i	$X_i \cdot f_i$	$X_i^2 \cdot f_i$
1	0	8	0	0
2	1	9	9	9
3	2	5	10	20
4	3	2	6	18
5	4	0	0	0
6	5	0	0	0
7	6	1	6	36
	total	25	31	83

$$s_X^2 = \frac{83 - (31)^2 / 25}{25 - 1} = 1{,}86$$

Variância Amostral

Considerando distribuição de freqüências de classes

$$S_X^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2 \cdot f_i}{n - 1}$$

equivalente ⇨

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

Exemplo da Fundição Y: diâmetro do furo (mm)

	Y_i	f_i	$Y_i \cdot f_i$	$Y_i^2 \cdot f_i$
1	11,77	5	58,85	692,66
2	11,90	3	35,7	424,83
3	12,03	7	84,21	1013,05
4	12,16	6	72,96	887,19
5	12,29	4	49,16	604,18
	total	25	300,88	3621,91

$$s_Y^2 = \frac{3621{,}91 - {(300{,}88)^2}/{25}}{25 - 1} = 0{,}032$$

Desvio Padrão DP(X)

População: $$\boxed{\sigma_X = \sqrt{\sigma_X^2}}$$

Amostra: $$\boxed{s_x \neq \sqrt{s_x^2}}$$

Empiricamente: $$\boxed{s_x = \frac{1}{c} \cdot \sqrt{s_x^2}}$$

onde c:

No Cap. 9 – Estimação de Parâmetros por Ponto será apresentada a justificativa da divisão por **c**

n	2	3	4	5	7	10	12	15	20	25	50	100	>100
c	0,7979	0,8862	0,9213	0,9400	0,9594	0,9727	0,9776	0,9823	0,9869	0,9896	0,9949	0,9975	1,0000

Exemplo da Fundição n = 25 c = 0,9896

- **X: número de defeitos por peça**

$$s_X^2 = 1{,}86$$

$$\boxed{s_X = \frac{1}{0{,}9896}\sqrt{1{,}86} = 1{,}38}$$

- **Y: diâmetro do furo (mm)**

$$s_Y^2 = 0{,}032$$

$$\boxed{s_Y = \frac{1}{0{,}9896}\sqrt{0{,}032} = 0{,}181}$$

<u>Dica "Excell":</u>

Selecionar: "fx" >> "Estatística" >> DESVPAD" Selecionar: células com a tabela de dados Clicar: "OK"

Obs.: Para População usar "DESVPADP"

Coeficiente de Variação $CV(X)$

Idéia: relação entre Desvio padrão e Média (%)

População:

$$\boxed{CV(X) = \frac{\sigma_X}{\mu_X} \cdot 100}$$

Amostra:

$$\boxed{CV(X) = \frac{s_x}{\bar{x}} \cdot 100}$$

Regra empírica:

CV < 5% ⇨ dispersão baixa

5% < CV < 15% ⇨ dispersão moderada

15% < CV < 30% ⇨ dispersão moderada

30% < CV < 50% ⇨ dispersão alta

CV > 50% ⇨ dispersão muito alta

Exemplo da Fundição:

X: número de defeitos por peça

$$cv(X) = 1{,}38 / 1{,}24 = 111{,}3\,\% \rightarrow \text{dispersão muito alta}$$

Y: diâmetro do furo

$$cv(Y) = 0{,}181 / 12{,}04 = 1{,}5\,\% \rightarrow \text{dispersão muito baixa}$$

Amplitude: R(X)

$$\boxed{R(X) = X_{máx} - X_{min}}$$

Exemplo da Fundição:

X: número de defeitos por peça

$X_{máx} = 6$

$X_{mín} = 0$

$R(X) = 6 - 0 = 6$

Y: diâmetro do furo (mm)

$Y_{máx} = 12,34$

$Y_{mín} = 11,73$

$R(Y) = 12,34 - 11,73 = 0,61$

Relação Empírica (útil para verificação de erros grosseiros):

$$\boxed{R/6 < s < R/3}$$

Exemplo da Fundição:

X: número de defeitos por peça $R(X) = 6$ $S_X = 1,38$

$6/6 < s < 6/3$ ⇨ $1 < S_x < 2$ ⇨ **OK!**

Y: diâmetro do furo (mm) $R(Y) = 0,61$ $S_Y = 0,181$

$0,61/6 < s < 0,61/3$ ⇨ $0,102 < s < 0,203$ ⇨ **OK!**

Momentos de Ordem “ t ” Centrado

$$\boxed{M_t = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^t}{n}}$$

Logo:

$$\boxed{M_1 = 0}$$

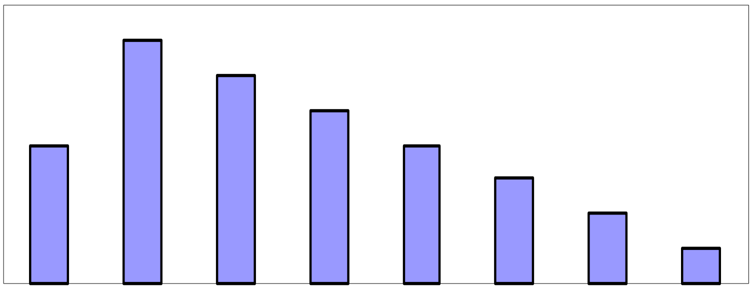
$$\boxed{M_2 = \frac{\sum_{1=}^{n} (X_i - \bar{X})^2}{n} = S^2 \cdot (\frac{n-1}{n})}$$

$$\boxed{M_3 = \frac{\sum_{i=}^{n} (X_i - \bar{X})^3}{n} = \frac{\sum_{i=1}^{n} x_i^3}{n} - 3 \cdot \bar{x} \cdot \frac{\sum_{i=1}^{n} x_i^2}{n} + 2 \cdot \bar{x}^3}$$

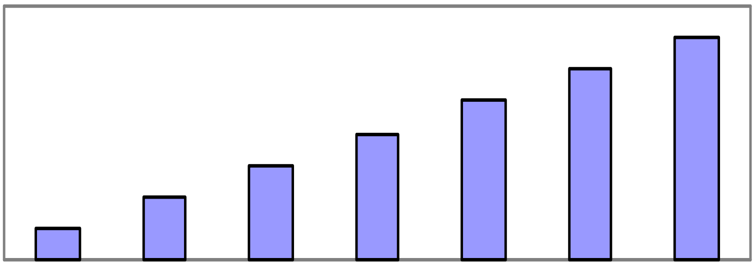
$$\boxed{M_4 = \frac{\sum (X_i - \bar{X})^4}{n} = \frac{\sum_{i=1}^{n} x_i^4}{n} - 4 \cdot \bar{x} \cdot \frac{\sum_{i=1}^{n} x_i^3}{n} + 6 \cdot \frac{\sum_{i=1}^{n} x_i^2}{n} - 3\bar{x}^4}$$

MEDIDAS DE ASSIMETRIA

Idéia: representa o grau de afastamento da condição de simetria



Assimetria Positiva



Assimetria Negativa

Coeficiente de Assimetria de Fisher (g_1):

$$\boxed{g_1 = \frac{n^2 \cdot M_3}{(n-1)\cdot(n-2)\cdot S^3}}$$

$g_1 < 0$: Assimetria Negativa
$g_1 = 0$: Simetria
$g_1 > 0$: Assimetria Positiva

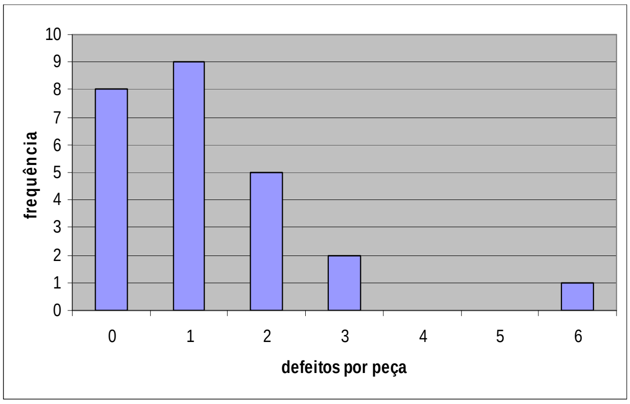
MEDIDAS DE ASSIMETRIA

Coeficiente de Assimetria de Fisher (g_1):

$$g_1 = \frac{n^2 \cdot M_3}{(n-1)\cdot(n-2)\cdot S^3}$$

Exemplo da Fundição:

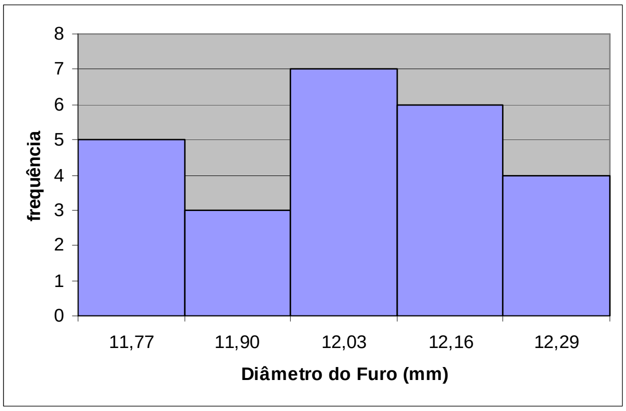
X: número de defeitos por peça



$$M_3(X) = \frac{\sum_{i=}^{k}(X_i - \overline{X})^3 \cdot f_i}{n} = 4{,}36$$

$$g_1 = \frac{25^2 \cdot 4{,}36}{(25-1)\cdot(25-2)\cdot 1{,}38^3} = 1{,}88$$

Y: diâmetro do furo (mm)

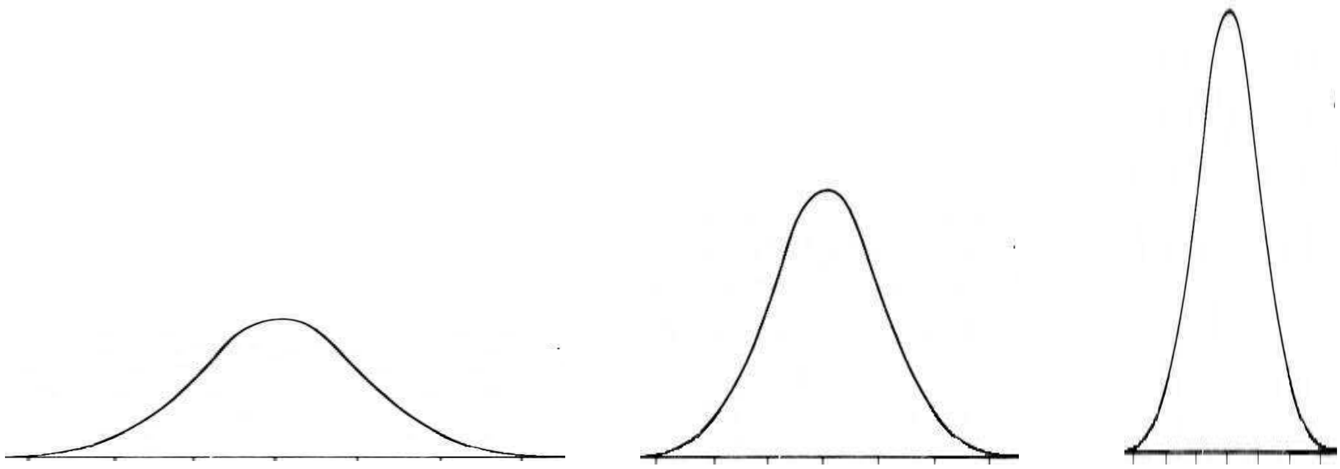


$$M_3(Y) = \frac{\sum_{1=1}^{k}(X_i - \overline{X})^3 \cdot f_i}{n} = 0{,}0009$$

$$g_1 = \frac{25^2 \cdot 0{,}009}{(25-1)\cdot(25-2)\cdot 0{,}181^3} = 0{,}172$$

Medidas de Achatamento ou Curtose

Idéia: representa o grau de achatamento comparado com a Gaussiana (distribuição Normal)



Curva Platicúrtica

Curva Mesocúrtica (Normal)

Curva Leptocúrtica

Coeficiente de Achatamento de Fisher (g_2):

$$\boxed{g_2 = \frac{n^2 \cdot (n+1) \cdot M_4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot S^4} - 3 \cdot \frac{(n-1)^2}{(n-2)\cdot(n-3)}}$$

$g_2 < 0 \rightarrow$ Curva Platicúrtica

$g_2 = 0 \rightarrow$ Curva Mesocúrtica (Normal)

$g_2 > 0 \rightarrow$ Curva Leptocúrtica

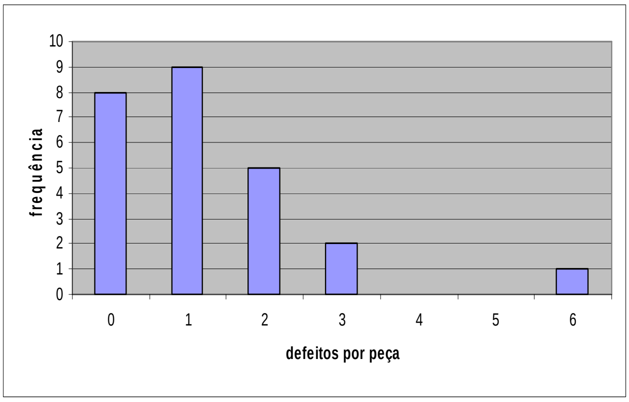
Medidas de Achatamento ou Curtose

Coeficiente de Achatamento de Fisher (g_2):

$$g_2 = \frac{n^2 \cdot (n+1) \cdot M_4}{(n-1)\cdot(n-2)\cdot(n-3)\cdot S^4} - 3 \cdot \frac{(n-1)^2}{(n-2)\cdot(n-3)}$$

Exemplo da Fundição:

X: número de defeitos por peça

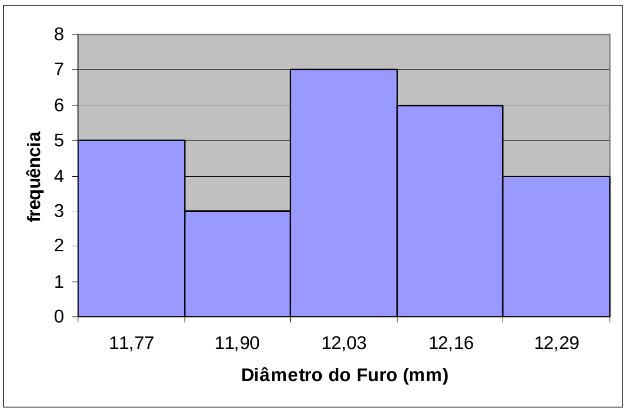


$$M_4(X) = \frac{\sum_{i=}^{k} (X_i - \overline{X})^4 \cdot f_i}{n} = 21{,}37$$

$$g_2(X) = 4{,}47$$

Curva Leptocúrtica

Y: diâmetro do furo (mm)



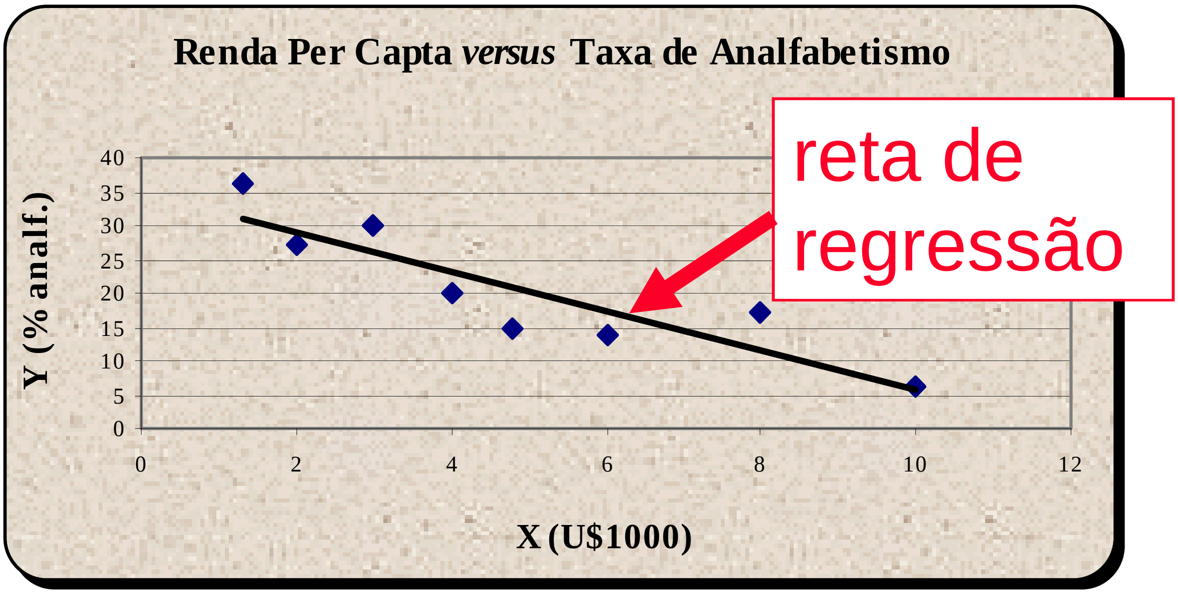
$$M_4(Y) = \frac{\sum_{1=1}^{k} (Y_i - \overline{Y})^4 \cdot f_i}{n} = 0{,}0017$$

$$g_2(Y) = -1{,}29$$

Curva Platicúrtica

CORRELAÇÂO LINEAR & RETA DE REGRESSÃO

Exemplo: amostra de 8 países
X : Renda Per Capita (U$ 1000)
Y : Taxa de Analfabetismo (%)



Verificação Visual:

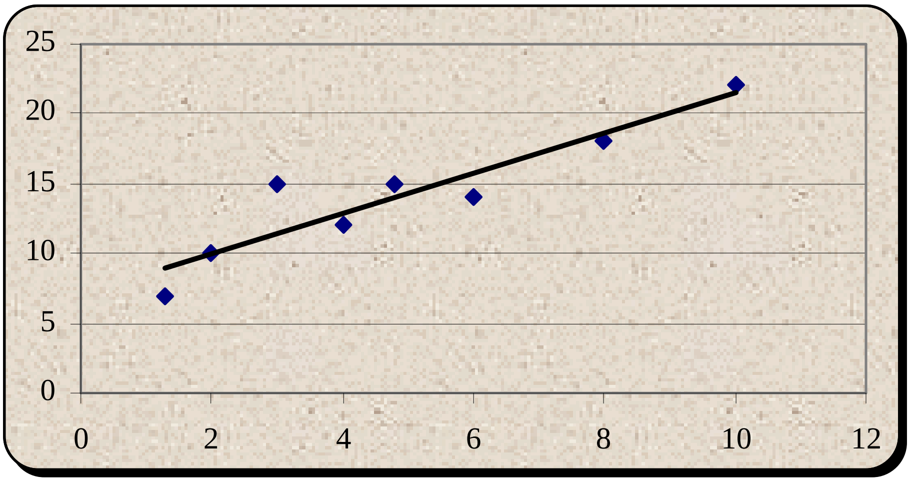
Existe tendência dos maiores valores de X corresponderem aos menores valores de Y, ou seja:

Existe *Correlação Linear Negativa* entre as variáveis

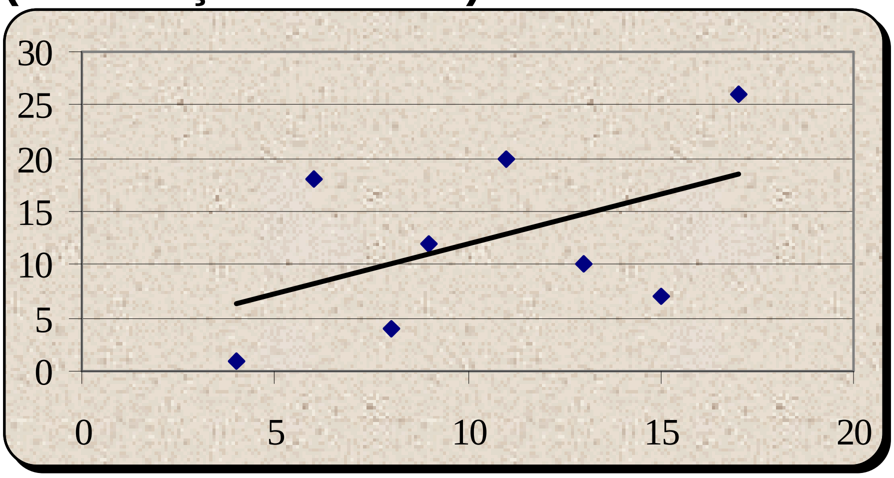
Uma ***Correlação Linear Positiva*** ocorre quando se verifica uma reta ascendente.

CORRELAÇÂO LINEAR

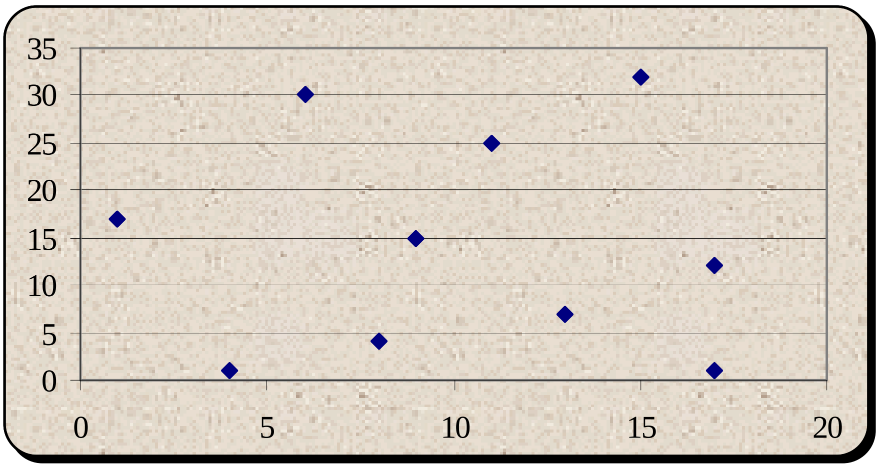
1) Grau Acentuado (Correlação Positiva):



2) Grau Moderado (Correlação Positiva):



3) Grau Nulo



MEDIDA DE CORRELAÇÂO LINEAR

Covariância: Mede a ***variabilidade*** considerando duas variáveis

$$S_{XY} = \operatorname{cov}(X,Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) * (y_i - \bar{y})}{n-1}$$

Caso particular:

$$S_{xx} = \operatorname{cov}(x,x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = S_x^2 = \quad \text{Variância Amostral}$$

Coeficiente de Correlação Linear de *Pearson*

População:

$$\rho = \frac{\operatorname{cov}(x,y)}{\sigma_x \cdot \sigma_y}$$

Amostra:

$$r = \frac{\operatorname{cov}(x,y)}{S_x \cdot S_y}$$

Dica “Excell”:

Selecionar: “fx” >> “Estatística” >> “Pearson

Selecionar: células com os dados da variável X

Selecionar: células com os dados da variável Y

Clicar: “OK”

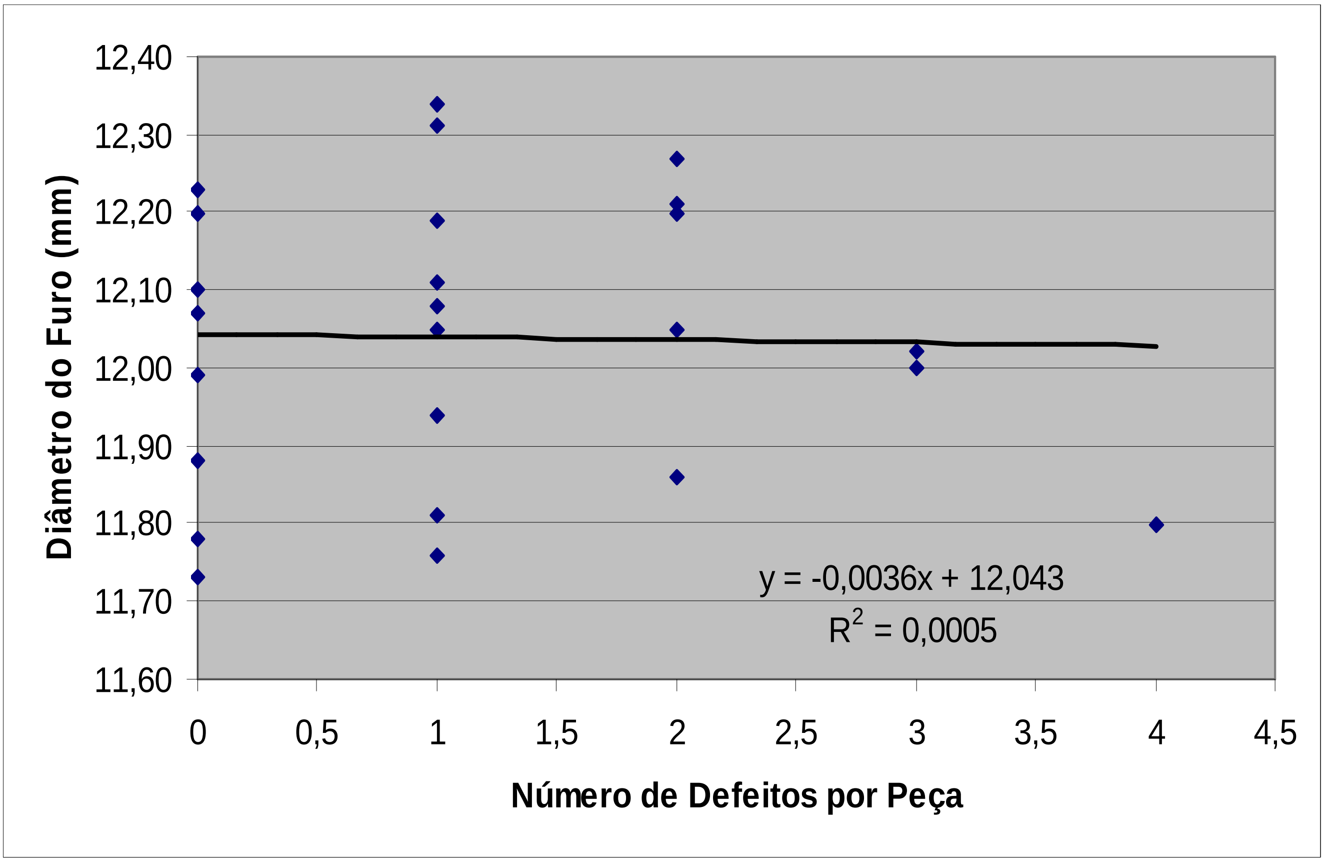
EXISTE CORRELAÇÃO ENTRE VARIÁVEIS ?

Variáveis:

D: Número de Defeitos por Peça

F: Diâmetro do Furo

DIAGRAMA DE DISPERSÃO E RETA DE REGRESSÃO



NÃO EXISTE EVIDÊNCIA DE CORRELAÇÃO ENTRE O DIÂMETRO DO FURO E O NÚMERO DE DEFEITOS POR PEÇA

CORRELAÇÃO: dicas usando Excell

Diagrama de Dispersão:

>> Selecionar: "Assistente de gráfico"

>> Tipo: Dispersão (XY)

>> Sub-tipo: só pontos;

>> Avançar;

>> Intervalo de dados: selecionar células com os dados da variável Y

>> Clicar na aba "Sequência

>> Valores de x: selecionar células com os dados da variável X

>> Avançar

>> Avançar

>> Concluir

Reta de Regressão:

>> Selecionar os pontos do gráfico gerado

>> Clicar aba "Gráfico" (painel superior)

>> Adicionar linha de tendência:

>> Tipo: linear

>> Clicar "opções": "exibir equação" e "exibir R-quadrado";

>> clicar "OK"