



PROBABILIDADE E ESTATÍSTICA

Profº Agnaldo Cieslak

Distribuição de frequências

A análise estatística de dados pode ocorrer por meio de dados agrupados e não agrupados.

Dados não agrupados são conjuntos ordenados e enumerados em quantidade menor que 30 elementos, não se aplicando a distribuição de frequência (DF).

Dados agrupados, formados por mais de 30 elementos, são chamados de grande conjunto de dados, onde se aplica a DF.

A DF é a primeira etapa para entender um problema:

- Método de agrupamento de dados em classes.
- Forma de resumir, organizar e apresentar os dados por meio de tabelas e gráficos, permitindo comparar conjuntos de dados.

Distribuição de frequências

Um dos objetivos da Distribuição de Frequência é diminuir o número de dados que serão trabalhados de modo direto, “modificando a forma de apresentação desses dados”. (Segundo Silva et al., p.18, 2006)

SILVA, Ermes Medeiros da et al. Estatística para os cursos de Economia, Administração e Ciências Contábeis. 3. ed. São Paulo: Atlas, 2006.

Através deste mecanismo, pode-se reorganizar os dados e agrupá-los de forma que o observador consiga identificar o que eles desejam transmitir, de forma facilitada.

Assim, vamos desenvolver passo a passo a elaboração de uma distribuição de frequências a partir de uma base de dados, observando a organização, cálculo e representação, demonstrando através de exemplo prático para melhor entendimento.

Dúvida em relação à construção pode ser tirada a partir do vídeo explicativo disponibilizado neste aula.

Por fim, os alunos devem resolver o exercício proposto da tarefa 5.

Distribuição de frequência

Em uma linha de produção de ações para navios, um estudo experimental sobre comprimento de trincas realizado em 50 chapas de aço foram observados os seguintes valores mostrados na tabela 1. O comprimento da trinca foi medido em milímetros. Considerando estes dados realizar:

a) Ordenar os dados;

b) Construir a distribuição de frequências;

a) Calcular h_t

b) Calcular k

c) Calcular h_c

d) Verificar $k \cdot h_c > h_t$

e) Definir limites de classe

f) Calcular o ponto médio de cada classe (X_i)

g) Contar os elementos de cada classe e estabelecer a frequência absoluta (f_i)

h) Calcular os tipos de frequência (%fr, fac, %frc, fad, %frd)

c) Interpretar os dados da tabela 2 e responder as questões:

a) Calcular o percentual de chapa de aço com comprimento de trinca inferior a 25mm

b) Calcular o percentual de chapa de aço com comprimento de trinca superior a 17mm.

c) Calcular o percentual de chapas de aço com comprimento da trinca superior a 26mm

d) Construir o histograma da distribuição de frequência do comprimento de trincas em chapas de aço.

e) Construir a ogiva de Galton para os valores de fac e fad

a) Interpretar a ogiva de Galton quanto ao resultado

Tabela 1 - comprimento de trincas em chapas de aço (mm)

15	12	13	17	20	22	19	21	18	27
7	9	14	22	27	26	13	8	6	28
24	17	12	21	25	28	23	20	10	14
23	21	16	15	5	26	21	18	19	22
29	27	19	13	6	9	20	19	14	13

Distribuição de frequência

Roteiro de trabalho para construção da distribuição de frequência:

1) Ordenar os dados do problema;

Tabela 1 ordenada- comprimento de trincas em chapas de aço (mm)

5	6	6	7	8	9	9	10	12	12
13	13	13	13	14	14	14	15	15	16
17	17	18	18	19	19	19	19	20	20
20	21	21	21	21	22	22	22	23	23
24	25	26	26	27	27	27	28	28	29

2) Construir a distribuição de frequências; (tabela 2)

a) Calcular ht (amplitude total dos dados)

$$ht = X_{\max} - X_{\min} = 29 - 5 = 24$$

b) Calcular k (número de classes) $n = \text{número de elementos da amostra} = 50$

$$k = 1 + 3,3 \log n \text{ (regra de Sturges)} = 1 + 3,3 \log 50 = 6,6066 \text{ classes} = \sim 7 \text{ classes}$$

c) Calcular hc (amplitude de intervalo de classes) Este valor é que define o intervalo de classes, pode ser inteiro ou não.

$$hc = ht/k = 24/7 = 3,42 \text{ (vamos arredondar para o inteiro maior)} = 4$$

d) Verificar $k \cdot hc > ht = 7 \cdot 4 > 24? \Rightarrow 28 > 24$ (ok, verificado)

Se o valor for menor que ht, então vc deve aumentar o valor de k ou hc até atender a condição.

Distribuição de frequência

e) Definir limites de classe

Classes são intervalos de variação da variável (neste caso a variável é: comprimento das trincas em mm).

Tabela 2:

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
1	5 ----- 9							
2	9 ----- 13							
3	13 ----- 17							
4	17 ----- 21							
5	21 ----- 25							
6	25 ----- 29							
7	29 ----- 33							
Σ								

Distribuição de frequência

f) Calcular o ponto médio de cada classe (X_i)

$X_i = (l_i + Li) / 2$ -> Li limite superior da classe e li é o limite inferior

g) Contar os elementos de cada classe e estabelecer a frequência absoluta (f_i)

Aqui apenas precisa contar os elementos de cada classe e colocar na tabela.

h) Calcular os tipos de frequência (%fr, fac, %frc, fad, %frd)

$\%fr = (f_i / \sum f_i) * 100$ -> representa a porcentagem de elementos em cada classe

$fac = \sum f_i (i=1, 2, 3, \dots, k)$ -> frequência absoluta acumulada crescente

$\%frc = (fac / \sum f_i) * 100$ -> representa a porcentagem acumulada crescente de elementos em cada classe

$fad = \sum f_i (i=k, \dots, 3, 2, 1)$ -> frequência absoluta acumulada decrescente. Obtemos somando a frequência de cada classe com todas frequências posteriores.

$\%frd = (fad / \sum f_i) * 100$ -> representa a porcentagem acumulada decrescente de elementos em cada classe

Distribuição de frequência

f) Calcular o ponto médio de cada classe (X_i)

$X_i = (li + Li) / 2$ → Li limite superior da classe e li é o limite inferior

Classes i	Comprimento mm li Li	X_i	f_i	%fr	fac	%frc	fad	%frd
1	5 ----- 9	7						
2	9 ----- 13	11						
3	13 ----- 17	15						
4	17 ----- 21	19						
5	21 ----- 25	23						
6	25 ----- 29	27						
7	29 ----- 33	31						
Σ								

Distribuição de frequência

g) Contar os elementos de cada classe e estabelecer a frequência absoluta (fi)

Aqui apenas precisa contar os elementos de cada classe e colocar na tabela.

Tabela 1 ordenada- comprimento de trincas em chapas de aço (mm)

5	6	6	7	8	9	9	10	12	12
13	13	13	13	14	14	14	15	15	16
17	17	18	18	19	19	19	19	20	20
20	21	21	21	21	22	22	22	23	23
24	25	26	26	27	27	27	28	28	29

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
1	5 ----- 9	7	5					
2	9 ----- 13	11	5					
3	13 ----- 17	15	10					
4	17 ----- 21	19	11					
5	21 ----- 25	23	10					
6	25 ----- 29	27	8					
7	29 ----- 33	31	1					
Σ			50					

Distribuição de frequência

h) Calcular os tipos de frequência (%fr, fac, %frc, fad, %frd)

→ $\%fr = (f_i / \sum f_i) * 100$ -> representa a porcentagem de elementos em cada classe

$fac = \sum f_i (i=1, 2, 3..., k)$ -> frequência absoluta acumulada crescente

$\%frc = (fac / \sum f_i) * 100$ -> representa a porcentagem acumulada crescente de elementos em cada classe

$fad = \sum f_i (i=k..., 3, 2, 1)$ -> frequência absoluta acumulada decrescente. Obtemos somando a frequência de cada classe com todas frequências posteriores.

$\%frd = (fad / \sum f_i) * 100$ -> representa a porcentagem acumulada decrescente de elementos em cada classe

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
1	5 ----- 9	7	5	10				
2	9 ----- 13	11	5	10				
3	13 ----- 17	15	10	20				
4	17 ----- 21	19	11	22				
5	21 ----- 25	23	10	20				
6	25 ----- 29	27	8	16				
7	29 ----- 33	31	1	2				
Σ			50					

Distribuição de frequência

h) Calcular os tipos de frequência (%fr, fac, %frc, fad, %frd)

%fr= $(f_i / \sum f_i) * 100$ -> representa a porcentagem de elementos em cada classe



fac= $\sum f_i$ (i=1, 2, 3...,k) -> frequência absoluta acumulada crescente

%frc= $(fac / \sum f_i) * 100$ -> representa a porcentagem acumulada crescente de elementos em cada classe

fad= $\sum f_i$ (i=k..., 3, 2, 1) -> frequência absoluta acumulada decrescente. Obtemos somando a frequência de cada classe com todas frequências posteriores.

%frd= $(fad / \sum f_i) * 100$ -> representa a porcentagem acumulada decrescente de elementos em cada classe

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
1	5 ----- 9	7	5	10	5			
2	9 ----- 13	11	5	10	10			
3	13 ----- 17	15	10	20	20			
4	17 ----- 21	19	11	22	31			
5	21 ----- 25	23	10	20	41			
6	25 ----- 29	27	8	16	49			
7	29 ----- 33	31	1	2	50			
Σ			50					

Distribuição de frequência

h) Calcular os tipos de frequência (%fr, fac, %frc, fad, %frd)

$\%fr = (fi / \sum fi) * 100$ -> representa a porcentagem de elementos em cada classe

$fac = \sum fi \ (i=1, 2, 3..., k)$ -> frequência absoluta acumulada crescente

➡ $\%frc = (fac / \sum fi) * 100$ -> representa a porcentagem acumulada crescente de elementos em cada classe

$fad = \sum fi \ (i=k..., 3, 2, 1)$ -> frequência absoluta acumulada decrescente. Obtemos somando a frequência de cada classe com todas frequências posteriores.

$\%frd = (fad / \sum fi) * 100$ -> representa a porcentagem acumulada decrescente de elementos em cada classe

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
1	5 ----- 9	7	5	10	5	10		
2	9 ----- 13	11	5	10	10	20		
3	13 ----- 17	15	10	20	20	40		
4	17 ----- 21	19	11	22	31	62		
5	21 ----- 25	23	10	20	41	82		
6	25 ----- 29	27	8	16	49	98		
7	29 ----- 33	31	1	2	50	100		
Σ			50					

Distribuição de frequência

h) Calcular os tipos de frequência (%fr, fac, %frc, fad, %frd)

$\%fr = (f_i / \sum f_i) * 100$ -> representa a porcentagem de elementos em cada classe

$fac = \sum f_i (i=1, 2, 3, \dots, k)$ -> frequência absoluta acumulada crescente

$\%frc = (fac / \sum f_i) * 100$ -> representa a porcentagem acumulada crescente de elementos em cada classe

➡ $fad = \sum f_i (i=k, \dots, 3, 2, 1)$ -> frequência absoluta acumulada decrescente. Obtemos somando a frequência de cada classe com todas frequências posteriores.

$\%frd = (fad / \sum f_i) * 100$ -> representa a porcentagem acumulada decrescente de elementos em cada classe

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
1	5 ----- 9	7	5	10	5	10	50	
2	9 ----- 13	11	5	10	10	20	45	
3	13 ----- 17	15	10	20	20	40	40	
4	17 ----- 21	19	11	22	31	62	30	
5	21 ----- 25	23	10	20	41	82	19	
6	25 ----- 29	27	8	16	49	98	9	
7	29 ----- 33	31	1	2	50	100	1	
Σ			50					

Distribuição de frequência


h) Calcular os tipos de frequência (%fr, fac, %frc, fad, %frd)

$\%fr = (fi / \sum fi) * 100$ -> representa a porcentagem de elementos em cada classe

$fac = \sum fi (i=1, 2, 3..., k)$ -> frequência absoluta acumulada crescente

$\%frc = (fac / \sum fi) * 100$ -> representa a porcentagem acumulada crescente de elementos em cada classe

$fad = \sum fi (i=k..., 3, 2, 1)$ -> frequência absoluta acumulada decrescente. Obtemos somando a frequência de cada classe com todas frequências posteriores.

 $\%frd = (fad / \sum fi) * 100$ -> representa a porcentagem acumulada decrescente de elementos em cada classe

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
1	5 ----- 9	7	5	10	5	10	50	100
2	9 ----- 13	11	5	10	10	20	45	90
3	13 ----- 17	15	10	20	20	40	40	80
4	17 ----- 21	19	11	22	31	62	30	60
5	21 ----- 25	23	10	20	41	82	19	38
6	25 ----- 29	27	8	16	49	98	9	18
7	29 ----- 33	31	1	2	50	100	1	2
Σ			50					

Distribuição de frequência

c) Interpretar os dados da tabela 2 e responder as questões:

a) Calcular o percentual de chapa de aço com comprimento de trinca inferior a 25mm

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
5	21 ----- 25	23	10	20	41	82	19	38

Vemos que 41 chapas de aço apresentam trincas com comprimento inferior a 25mm (fac).
Ou seja, 82% das chapas produzidas têm trincas menores que 25mm. (%frc)

b) Calcular o percentual de chapa de aço com comprimento de trinca superior a 17mm.

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
4	17 ----- 21	19	11	22	31	62	30	60

Olhando a 4ª classe da tabela2, vemos que 60% das chapas produzidas têm trincas maiores que 17mm.

Esta informação é útil para verificar a capacidade do processo ou produto atender a necessidade do mercado.

c) Calcular o percentual de chapas de aço com comprimento da trinca superior a 26mm

Aqui como o valor de 26mm está entre o limite inferior de 25mm e o limite superior de 29mm, não há como tirar diretamente da tabela. Precisamos usar a chamada interpolação linear do %frd.

Classes i	Comprimento mm li Li	Xi	fi	%fr	fac	%frc	fad	%frd
6	25 ----- 29	27	8	16	49	98	9	18

Distribuição de frequência

c) Interpretar os dados da tabela 2 e responder as questões:

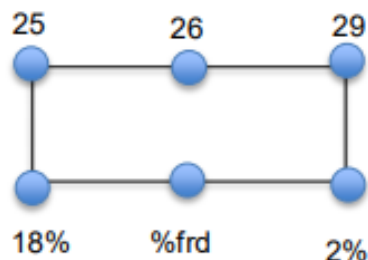
c) Calcular o percentual de chapas de aço com comprimento da trinca superior a 26mm

Aqui como o valor de 26mm está entre o limite inferior de 25mm e o limite superior de 29mm, não há como tirar diretamente da tabela. Precisamos usar a chamada interpolação linear do %frd.

Classes i	Comprimento mm li	Xi	fi	%fr	fac	%frc	fad	%frd
6	25 ----- 29	27	8	16	49	98	9	18
7	29 ----- 33	31	1	2	50	100	1	2

Vamos usar a classe 5 e 6 para fazer a interpolação linear.

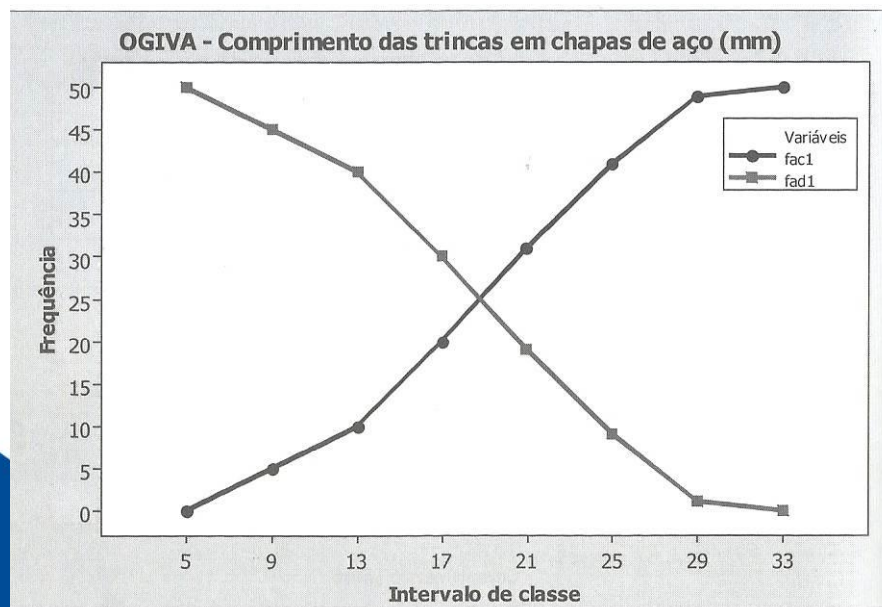
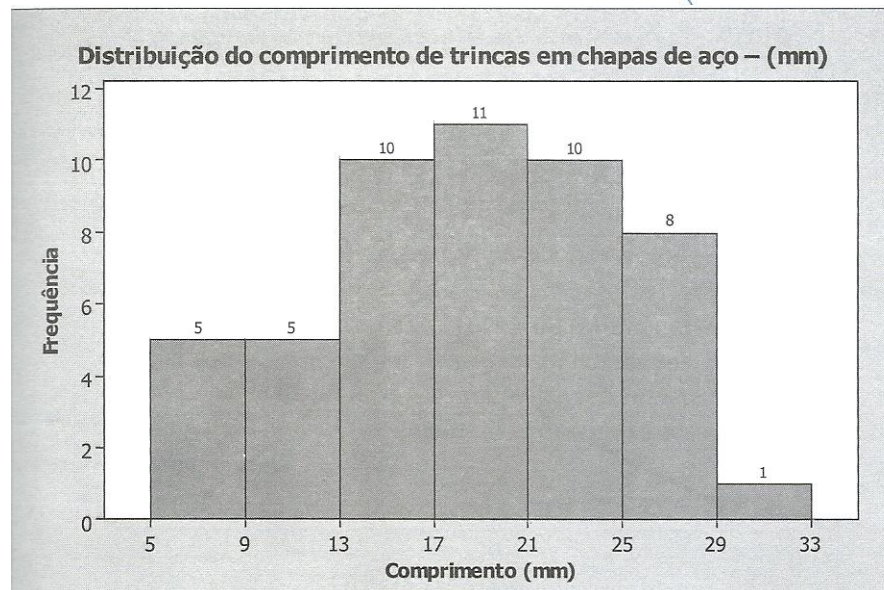
Aqui como o valor de 26mm está entre o limite inferior de 25mm e o limite superior de 29mm, não há como tirar diretamente da tabela. Precisamos usar a chamada interpolação linear do %frd.



$$\frac{29-25}{29-26} = \frac{2-18}{2-\%frd} \Rightarrow \frac{4}{3} = \frac{-16}{2-\%frd} \Rightarrow \%frd=14,00$$

Assim podemos dizer que 14% das chapas de aço produzidas apresentam comprimento de trinca acima de 26mm.

Conclui-se que 5 chapas das 50 têm trincas inferiores a 9mm, ou seja 10%.



Conclui-se pela linha fad1 que 45 de 50 chapas analisadas tem comprimento de trinca superior a 9mm, ou seja, 90% das chapas estariam reprovadas.