CONCEITOS INICIAIS DE ESTATÍSTICA

MÓDULO 2 – DISTRIBUIÇÃO DE FREQÜÊNCIA - ELEMENTOS

Prof. Rogério Rodrigues

I) TABELA PRIMITIVA E DISTRIBUIÇÃO DE FREQÜÊNCIA :

No processo de amostragem , a forma de registro mais imediata pode ser uma tabela primitiva , onde os dados coletados se apresentam sem qualquer ordenação sistemática , é o que chamamos de *tabela primitiva* ou *rol* .

Exemplo : Uma pesquisa sobre a idade dos 36 alunos de uma turma de Administração . Os dados serão simplesmente registrados em forma de tabela , até mesmo sem cabeçalho :

19	27	23	20	27	27	25	19	23
28	29	31	19	24	29	22	21	20
30	32	31	21	24	31	21	27	26
28	33	33	32	26	34	25	20	22

Essa tabulação primitiva poderia ser expressa em uma ordem mais conveniente , como , por exemplo , ordem crescente das idades :

19	20	21	23	25	27	28	31	32
19	20	22	24	26	27	29	31	33
19	21	22	24	26	27	29	31	33
20	21	23	25	27	28	30	32	34

Nessa última tabulação , algumas coisas ficam mais claras , como , por exemplo , o número de alunos com idades compreendidas dentro de determinado intervalo , alunos com idades entre 20 e 30 anos ou alunos com idades superiores a 25 anos . Mas se desejamos verificar a freqüência com que cada idade aparece , deveremos lançar mão de uma tabela como a seguinte , onde cada idade aparece assinalada com o número de vezes em que foi registrada , ou seja , a freqüência com que apareceu na amostragem .

Idade (anos)	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Freq.	3	3	3	2	2	2	2	2	4	2	2	1	3	2	2	1

Essa tabela ainda poderia ser apresentada por faixas de idade , ou classes , assim

FAIXA ETÁRIA (anos)	FREQÜÊNCIA
19 I– 23	11
23 I– 27	8
27 I– 31	9
31 I– 35	8

II) ELEMENTOS DE UMA DISTRIBUIÇÃO DE FREQÜÊNCIA :

II.1) CLASSES :

São faixas ou intervalos de variação da variável . No nosso exemplo anterior , são os intervalos registrados na primeira coluna da última tabela ; essas classes são genericamente assinaladas por $\mathbf{i_n}$, sendo **n** , o número da classe . Assim , no nosso exemplo temos

i_1 = primeira classe : 19 I– 23
i_2 = segunda classe : 23 I– 27
i_3 = terceira classe : 27 I– 31
i_4 = quarta classe : 31 I– 35

Cada classe $\mathbf{i_n}$ está associada a uma freqüência $\mathbf{f_n}$, registrada na segunda coluna de nossa última tabela ; assim temos

$f_1 = 11$, $f_2 = 8$, $f_3 = 9$, $f_4 = 8$ e $\sum_{n=1}^{4} f_n = 36$ (total de alunos da turma)

II.2) LIMITES DE CLASSE :

São os extremos de cada classe . Denotaremos o limite superior por L_i e o limite inferior por ℓ_i . Assim , no nosso exemplo , temos

na primeira classe : $L_1 = 23$ e $\ell_1 = 19$
na segunda classe : $L_2 = 27$ e $\ell_2 = 23$
na terceira classe : $L_3 = 31$ e $\ell_3 = 27$
na quarta classe : $L_4 = 35$ e $\ell_4 = 31$

II.3) AMPLITUDE DE UM INTERVALO DE CLASSE :

É a diferença entre os limites superior L_n e inferior ℓ_n . Então , temos $\boxed{h_i = L_i - \ell_i}$

No nosso exemplo , todas as classes têm a mesma amplitude e , então , podemos registrar que $\mathbf{h_1 = h_2 = h_3 = h_4} = 23 - 19 = 27 - 23 = 31 - 27 = 35 - 31 = 4$ anos .

II.4) AMPLITUDE TOTAL DA DISTRIBUIÇÃO :

É a diferença entre o limite superior da última classe $L_{máx}$ e o limite inferior da primeira classe ℓ_{min} ; ou seja , $\Delta T = L_{máx} - \ell_{min}$. No nosso exemplo , $\Delta T = 35 - 19 = 16$ anos .

II.5) AMPLITUDE AMOSTRAL :

É a diferença entre o valor máximo e o valor mínimo constantes na amostra , ou seja , $AA = x_{máx} - x_{mín}$. No presente caso , temos que $AA = 34 - 19 = 15$ anos .

II.6) PONTO MÉDIO DE UMA CLASSE :

È o ponto que divide uma classe em duas partes iguais , ou seja , é a média aritmética entre

os limites L_i e ℓ_i de uma classe . Então , podemos escrever $x_i = \frac{\ell_i + L_i}{2}$. No nosso exemplo

$x_1 = \frac{19+23}{2} = 21$ anos ,

$x_2 = \frac{23+27}{2} = 25$ anos ,

$x_3 = \frac{27+31}{2} = 29$ anos e

$x_4 = \frac{31+35}{2} = 33$ anos .

II.7) FREQÜÊNCIA RELATIVA

É a razão entre cada freqüência absoluta de uma classe e a freqüência total da distribuição , ou seja ,

$$fr_i = \frac{f_i}{\sum f_i}$$

No nosso exemplo , tem-se :

- $fr_1 = \frac{11}{36} = 0{,}306$ ou 30,6 %
- $fr_2 = \frac{8}{36} = 0{,}222$ ou 22,2 %
- $fr_3 = \frac{9}{36} = 0{,}25$ ou 25%
- $fr_4 = \frac{8}{36} = 0{,}222$ ou 22,2% .

Observe que $\sum fr_i = 100\% = 1$.

II.8) FREQÜÊNCIA ACUMULADA :

a) Freqüência Acumulada "abaixo de "

É o total das freqüências de todos os valores inferiores ao limite superior de uma determinada classe ; assim , temos , genericamente , $F_i = f_1 + f_2 + f_3 + \ ... \ f_{i-2} + f_{i-1} + f_i$ ou ainda

$$F_i = \sum_{k=1}^{i} f_k$$

No nosso exemplo inicial temos :

→ $F_1 = f_1 = 11$ (Número de alunos com idade inferior a 23 anos)
→ $F_2 = f_1 + f_2 = 11 + 8 = 19$ (Número de alunos com idade inferior a 27 anos)
→ $F_3 = f_1 + f_2 + f_3 = 11 + 8 + 9 = 28$ (Número de alunos com idade inferior a 31 anos)
→ $F_4 = f_1 + f_2 + f_3 + f_4 = 11 + 8 + 9 + 8 = 36$ (Total de alunos da pesquisa)

b) Freqüência Acumulada "acima de "

É o total das freqüências de todos os valores iguais ou superiores ao limite superior de uma determinada classe ; assim , numa distribuição com n classes , temos , genericamente

$$F_i = f_i + f_{i+1} + f_{i+2} + \ ... + f_{n-1} + f_n \qquad \text{ou} \qquad F_i = \sum_{k=i}^{n} f_k$$

No nosso exemplo inicial , temos

→ $F_1 = 11 + 8 + 9 + 8 = 36$ (Alunos com idade maior ou igual a 19 anos)

→ $F_2 = 8 + 9 + 8 = 25$ (Alunos com idade maior ou igual a 23 anos)

→ $F_3 = 9 + 8 = 17$ (Alunos com idade maior ou igual a 27 anos)

→ $F_4 = 8$ (Alunos com idade maior ou igual a 31 anos)

II.9) FREQÜÊNCIA ACUMULADA RELATIVA DE UMA CLASSE :

É a razão entre a freqüência acumulada da classe e a freqüência total da distribuição ; ou seja

$$Fr_i = \frac{F_i}{\sum f_i}$$

Do nosso exemplo inicial , podemos calcular , por exemplo

a) Percentual dos alunos com idade inferior a 31 anos = $\frac{28}{36} = 0{,}78 = 78\%$ ("para baixo")

b) Percentual dos alunos com idade superior ou igual a 23 anos = $\frac{25}{26} = 0{,}69 = 69\%$ ("para cima")

II.10) NÚMERO DE CLASSES NUMA DISTRIBUIÇÃO:

Uma das regras mais utilizadas para se determinar o número de classes i em função do número total de dados n é a *Regra de Sturges*; essa regra é dada pela fórmula

$$i \equiv 1 + 3{,}3.\log n$$

Em que n é o número total de dados e i é o número recomendado de classes.

Exemplo : Se o número total de dados é 40 , teremos $i \equiv 1 + 3{,}3.\log 40 \equiv 1 + 3.3.(1{,}602)$ e teremos, finalmente $i \equiv 6{,}28$, ou seja, 6 classes.

A aplicação desse fórmula para alguns valores de n gera a seguinte tabela, de i em função de n:

n	3 a 5	6 a 11	12 a 22	23 a 46	47 a 90	91 a 181
i	3	4	5	6	7	8

EXERCÍCIOS PROPOSTOS :

Em todos os exercícios , considere a seguinte notação :

i = número de classes numa distribuição de freqüência ;
ℓ_i = limite inferior de uma classe numa distribuição de freqüência;
L_i = limite superior de uma classe numa distribuição de freqüência;
h_i = amplitude de um intervalo de classe ;
ΔT = amplitude total da distribuição de freqüência;
AA = amplitude amostral ;
x_i = ponto médio de uma classe numa distribuição de freqüência ;
f_i = freqüência absoluta de uma classe ;
fr_i = freqüência relativa de uma classe ;
F_i = freqüência acumulada ;
Fr_i = freqüência acumulada relativa .

1) Em cada tabulação descrita a seguir , faça o que se pede :
1º) Converta a tabela dada numa tabela ordenada ;
2º) Construa uma distribuição de freqüências de classes com a mesma amplitude e
a) calcule a amplitude das classes ,
b) calcule a amplitude total da distribuição ,
c) calcule a amplitude amostral ,
d) calcule o ponto médio de cada classe .
3º) Construa e preencha uma tabela de freqüências com o formato abaixo :

i	classes	f_i	x_i	fr_i	F_i	Fr_i (para baixo)	Fr_i (para cima)

Tabela 1 : Salários , em reais , de 30 funcionários de uma empresa

240	280	400	250	260	520
320	240	490	370	500	240
530	450	370	280	270	280
260	270	440	450	490	510
510	420	500	240	315	450

Tabela 2 : Estatura , em centímetros , de 24 atletas de um clube .

202	174	193	210	167	184
193	208	175	200	175	191
200	166	181	163	181	205
165	200	190	168	205	197

Tabela 3 : Número de ações de uma mesma empresa pertencentes a 36 acionistas .

250	125	90	320	84	240	310	500	450
78	85	177	290	215	510	338	180	275
500	279	357	310	336	378	400	225	520
96	150	535	496	190	150	230	480	425

Tabela 4 : Notas de 45 candidatos numa prova de seleção para um emprego .

12	4	12	2	15	3	9	15	15
18	5	0	2	13	7	9	0	10
6	10	15	16	19	5	4	19	1
12	18	10	17	8	16	17	1	4
20	4	2	15	15	11	8	14	12

2) LIVRO : Crespo , Antônio Arnot , Estatística Fácil , São Paulo , Ed. Saraiva , 2.002 .
Exercícios das páginas 62 , 66 , 67 , 68 e 69 .