**DISTRIBUIÇÃO DE FREQUÊNCIA**

 É um tipo de tabela **que condensa uma coleção de dados conforme as freqüências** (repetições de seus valores).

***Tabela primitiva ou dados brutos:*** É uma **tabela ou relação de elementos que não foram numericamente organizados**. É difícil formarmos uma idéia exata do comportamento do grupo como um todo, a partir de dados não ordenados.

*Ex : 45, 41, 42, 41, 42 43, 44, 41 ,50, 46, 50, 46, 60, 54, 52, 58, 57, 58, 60, 51*

***ROL:*** É a **tabela obtida após a ordenação dos dados** (crescente ou decrescente).

*Ex : 41, 41, 41, 42, 42 43, 44, 45 ,46, 46, 50, 50, 51, 52, 54, 57, 58, 58, 60, 60*

***Distribuição de frequência sem intervalos de classe:*** É a simples condensação dos dados conforme as repetições de seus valores. Para um **ROL** de tamanho razoável esta distribuição de frequência é inconveniente, já que exige muito espaço. Veja exemplo abaixo:

|  |  |
| --- | --- |
| Dados | Frequência |
| 41 | 3 |
| 42 | 2 |
| 43 | 1 |
| 44 | 1 |
| 45 | 1 |
| 46 | 2 |
| 50 | 2 |
| 51 | 1 |
| 52 | 1 |
| 54 | 1 |
| 57 | 1 |
| 58 | 2 |
| 60 | 2 |
| Total | 20 |

***Distribuição de frequência com intervalos de classe:*** Quando o tamanho da amostra é elevado, é mais racional efetuar o agrupamento dos valores em vários intervalos de classe.

60-41 = 19 $\sqrt{n}= \sqrt{20}=4,47$ (**linhas da tabela**)

h= (de quanto em quanto varia o intervalo) h = $\frac{19}{5}=3,8 $arredonda 4

**Ou Regra de Sturges = k = 1 +3,3. log (n) ou** $\sqrt{n}=$

K = 1 + 3,3 . log 20

K = 1+ 3,3 . 1,3

K= 1 + 4,29 = 5,3 arredonda **5 linhas**

|  |  |
| --- | --- |
| Classes | Frequências |
| 41 |------- 45 | 7 |
| 45 |------- 49 | 3 |
| 49 |------- 53 | 4 |
| 53 |------- 57 | 1 |
| 57 |------- 61 | 5 |
| Total | 20 |

**ELEMENTOS DE UMA DISTRIBUIÇÃO DE FREQÜÊNCIA (com intervalos de classe) **

**CLASSE:** são os intervalos de variação da variável e é simbolizada por ***i*** e o número total de classes simbolizada por ***k***. Ex: na tabela anterior ***k*** = 5 e **49 |------- 53** é a **3ª classe**, onde ***i*** = 3.

**LIMITES DE CLASSE:** são os extremos de cada classe. O menor número é o limite inferior de classe ( **li** ) e o maior número, limite superior de classe ( **Li** ). *Ex: em* ***49 |------- 53,... l3*** *= 49 e* ***L3*** *= 53. O símbolo* ***|-------*** *representa* ***um intervalo fechado à esquerda e aberto à direita.*** *O dado 53 do* ***ROL*** *não pertence a classe 3 e sim a classe 4 representada por* ***53 |------- 57****.*

**AMPLITUDE DO INTERVALO DE CLASSE:** é obtida através da diferença entre o limite superior e inferior da classe e é simbolizada por ***hi = Li - li***. *Ex: na tabela anterior* ***hi*** *= 53 - 49 = 4.* ***Obs****: Na distribuição de freqüência c/ classe o* ***hi será igual em todas as classes.***

**AMPLITUDE TOTAL DA DISTRIBUIÇÃO:**  é a diferença entre o limite superior da última classe e o limite inferior da primeira classe. ***AT = L(max) - l(min)***. *Ex: na tabela anterior* ***AT*** *= 61 - 41= 20.*

**AMPLITUDE TOTAL DA AMOSTRA (ROL):** é a diferença entre o valor máximo e o valor mínimo da amostra (ROL). Onde ***AA = Xmax - Xmin***. Em nosso exemplo ***AA*** = 60 - 41 = 19.

**Obs**: ***AT*** sempre será maior que ***AA***.

**PONTO MÉDIO DE CLASSE**: é o ponto que divide o intervalo de classe em duas partes iguais. *.......Ex: em* ***49 |------- 53*** *o ponto médio* ***x3*** *= (53+49)/2 = 51, ou seja* ***x3=( l3 + L3 )/2.***

**Método prático para construção de uma Distribuição de Freqüências c/ Classe **

**1º -** Organize os dados brutos em um **ROL**.

**2º -** Calcule a **amplitude amostral** ***AA***.

* + - No nosso exemplo: ***AA*** = 60 - 41 = 19

**3º -** Calcule o número de classes através da "**Regra de Sturges": ou**

 Nesta expressão:

– k é o número de classes.

– N é o número total de observações na amostra.

– Log é o logaritmo comum da base 10.

**Obs:** Qualquer regra para determinação do nº de classes da tabela não nos levam a uma decisão final; esta vai depender, na realidade de um julgamento pessoal, que deve estar ligado à natureza dos dados.

No nosso exemplo: ***n*** = 20 dados, então ,a princípio, a regra sugere a adoção de 5 classes.

**4º -** Decidido o nº de classes, calcule então a amplitude do intervalo de classe**h > AA / i**.

No nosso exemplo: **AA/i** = 19/5 = 3,8 . Obs: Como **h > AA/i** um valor ligeiramente superior para haver folga na última classe. Utilizaremos então ***h*** = 4

**5º -** Temos então o menor nº da amostra, o nº de classes e a amplitude do intervalo. Podemos montar a tabela, com o cuidado para não aparecer classes com freqüência = 0 (zero).

No nosso exemplo: o menor nº da amostra = 41 + ***h*** = 45, logo a primeira classe será representada por ...... **41 |------- 45**. As classes seguintes respeitarão o mesmo procedimento.

O primeiro elemento das classes seguintes sempre serão formadas pelo último elemento da classe anterior.

**TIPOS DE FREQUÊNCIAS**

***Freqüências simples ou absoluta:*** são os valores que realmente representam o número de dados de cada classe. A soma das freqüências simples é igual ao número total dos dados da distribuição.

***Freqüências relativas:*** são os valores das razões entre as freqüência absolutas de cada classe e a freqüência total da distribuição. A soma das **freqüências relativas é igual a 1 (100 %).**

.

**Polígono de freqüência:** é um gráfico em linha, sendo as freqüências marcadas sobre perpendiculares ao eixo horizontal, levantadas pelos pontos médios dos intervalos de classe. Para realmente obtermos um polígono (linha fechada), devemos completar a figura, ligando os extremos da linha obtida aos pontos médios da classe anterior à primeira e da posterior à última, da distribuição.

.

**Polígono de freqüência acumulada:** é traçado marcando-se as freqüências acumuladas sobre perpendiculares ao eixo horizontal, levantadas nos pontos correspondentes aos limites superiores dos intervalos de classe.

***Freqüência simples acumulada de uma classe:*** é o total das freqüências de todos os valores inferiores ao limite superior do intervalo de uma determinada classe.

***Freqüência relativa acumulada de uma classe:*** é a freqüência acumulada da classe, dividida pela frequência total da distribuição.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **CLASSE** | **f freq. simples** | **xi Ponto Médio** | **fr freq. relativa** | **Fr Percentual** | **FA** | **FAr** |
| 50 |-------- 54  | 4 |  $\frac{50+54}{2}$ 52 | $\frac{4}{40}= $0,100 | 10% | 4 | 0,100 |
| 54 |-------- 58 | 9 | 56 | $\frac{9}{40}= $0,225 | 22,5% | 13 | 0,325 |
| 58 |-------- 62 | 11 | 60 | $\frac{11}{40}= $0,275 | 27,5% | 24 | 0,600 |
| 62 |-------- 66 | 8 | 64 | $\frac{8}{40}= $0,200 | 20% | 32 | 0,800 |
| 66 |-------- 70 | 5 | 68 | $\frac{5}{40}= $0,125 | 12,5% | 37 | 0,925 |
| 70 |-------- 74 | 3 | 72 | $\frac{3}{40}= $0,075 | 7,5% | 40 | 1,000 |
| Total | 40 |  | 1,000 | 100% |  |  |

**f** = frequência simples; **xi** = ponto médio de classe; **fr** = frequência relativa

**FA**= frequência simples acumulada; e **FAr** = frequência relativa acumulada.

* **Obs:** uma distribuição de freqüência sem intervalos de classe é representada graficamente por um diagrama onde cada valor da variável é representado por um segmento de reta vertical e de comprimento proporcional à respectiva freqüência.

.

**EXERCÍCIO**

Os dados abaixo referem-se ao número de pessoas que residem em uma amostra de 35 domicílios do bairro Vila Nova

2 3 4 4 5 3 4

5 6 5 3 1 5 5

1 3 4 5 5 5 3

2 2 5 4 4 2 3

5 4 5 4 2 4 9

a) Construa uma distribuição de frequência em classes.

b) Determine as frequências simples acumuladas de cada classe.

c) Determine as frequências relativas de cada classe.

d) Qual o percentual de pessoas que residem mais do que 4 em um domicílio?

e) Qual o percentual de pessoas que residem até duas pessoas em um domicílio?

.