Revisão dos Conceitos Básicos de Química e Estatística

Revisão dos Conceitos Básicos de Estatística

- Definições Básicas de Estatística
- Média
- Separatrizes (Quartil, Decis e Percentil)
- Desvio Padrão
- Variância
- Função de Distribuição de Probabilidade
- Tamanho da Amostra

Definições Básicas da Estatística

- **FENÔMENO ESTATÍSTICO**: é qualquer evento que se pretenda analisar, cujo estudo seja possível da aplicação do método estatístico. São divididos em três grupos:
  - **Fenômenos de massa ou coletivo**: são aqueles que não podem ser definidos por uma simples observação. A estatística dedica-se ao estudo desses fenômenos. Ex: A natalidade na Grande Vitória, O preço médio da cerveja no Espírito Santo, etc.
  - **Fenômenos individuais**:são aqueles que irão compor os fenômenos de massa. Ex: cada nascimento na Grande Vitória, cada preço de cerveja no Espírito Santo, etc.
  - **Fenômenos de multidão**:quando a s características observadas para a massa não se verificam para o particular.

Definições Básicas da Estatística

- **DADO ESTATÍSTICO**: é um dado numérico e é considerado a matéria-prima sobre a qual iremos aplicar os métodos estatísticos.
- **POPULAÇÃO**: é o conjunto total de elementos portadores de, pelo menos, uma característica comum.
- **AMOSTRA**: é uma parcela representativa da população que é examinada com o propósito de tirarmos conclusões sobre a essa população.
- **PARÂMETROS**: São valores singulares que existem na população e que servem para caracterizá-la.Para definirmos um parâmetro devemos examinar toda a população. Ex: Os alunos deste curso têm em média 1,70 metros de estatura.
- **ESTIMATIVA**: é um valor aproximado do parâmetro e é calculado com o uso da amostra.
- **ATRIBUTO**: quando os dados estatísticos apresentam um caráter qualitativo, o levantamento e os estudosnecessários ao tratamento desses dados são designados genericamente de estatística de atributo.

Definições Básicas da Estatística

- **VARIÁVEL**: É, convencionalmente, o conjunto de resutados possíveis de um fenômeno.
  - **VARIÁVEL QUALITATIVA**: Quando seu valores são expressos por atributos: sexo, cor da pele,etc.
  - **VARIÁVEL QUANTITATIVA**: Quando os dados são de caráter nitidamente quantitativo, e o conjunto dos resultados possui uma estrutura numérica, trata-se portanto da estatística de variável.

Definições Básicas da Estatística

- **VARIÁVEIS QUANTITATIVAS**:
  - **VARIÁVEL DISCRETA OU DESCONTÍNUA**: Seus valores são expressos geralmente através de números inteiros não negativos. Resulta normalmente de contagens. Ex: Nº de alunos presentes às aulas de Qualidade Ambiental Química no 1º semestre de 1997: mar = 40 , abr = 30 , mai = 35 , jun = 36.
  - **VARIÁVEL CONTÍNUA**: Resulta normalmente de uma mensuração, e a escala numérica de seus possíveis valores corresponde ao conjunto R dos números Reais, ou seja, podem assumir, teoricamente, qualquer valor

Definições Básicas da Estatística

DISTRIBUIÇÃO DE FREQUÊNCIA

- É um tipo de tabela que condensa uma coleção de dados conforme as frequências (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

Definições Básicas da Estatística

- **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 seu valores. Para um ROL de tamanho razoável esta distribuição de frequência é inconveniente, já que exige muito espaço. Veja exemplo ao lado:

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

Definições Básicas da Estatística

- **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.

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

Medidas de Posição ou Médias

- São as estatísticas que representam uma série de dados orientando-nos quanto à posição da distribuição em relação ao eixo horizontal do gráfico da curva de frequência.
- As medidas de posições mais importantes são as medidas de tendência central ou promédias (verifica-se uma tendência dos dados observados a se agruparem em torno dos valores centrais).
- As medidas de tendência central mais utilizadas são: **média aritmética** e a mediana. Outros menos usados são as médias: **geométrica**, **harmônica**, quadrática, cúbica e biquadrática. As outras medidas de posição são as separatrizes, que englobam: os decis, os quartis e os percentis.

Média Aritimética

- É igual ao quociente entre a soma dos valores do conjunto e o número total dos valores.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

média

onde x_i são os valores da variável

n o número de valores.

Média Aritimética

- Exemplo: Sabendo-se que a concentração de NOx medida em uma região, durante o período de 1 hora em intervalos de 10 em 10 minutos, foi de 10, 14, 13, 15, 16 e 18 μg/m^3. Assim, temos uma concentração média horária de:

valores da variável

$$\bar{x} = \frac{\overline{(10+14+13+15+16+18)}}{6} = 14 \mu g / m^3$$

média

número de valores.

Média Aritimética

- Desvio em relação à média: é a diferença entre cada elemento de um conjunto de valores e a média aritmética, ou seja:

$$d_i = x_i - \bar{x}_i$$

- No exemplo anterior temos seis desvios:
  - $d_1 = 10 - 14 = -4$
  - $d_2 = 14 - 14 = 0$
  - $d_3 = 13 - 14 = -1$
  - $d_4 = 15 - 14 = 1$
  - $d_5 = 16 - 14 = 2$
  - $d_6 = 18 - 14 = 4$

Média Geométrica

- É a raiz n-ésima do produto de todas as variáveis

n o número de variáveis

$$\bar{x}_g = \sqrt[n]{x_1 . x_2 . x_3 . x_4 . x_5 \ldots . x_n}$$

média

onde x_i são os valores da variável

Média Geométrica

Exemplo

- Calcular a média geométrica dos seguintes conjuntos de números:
  - a) { 10, 60, 360 }........ no excel : =(10*60*360)^(1/3) ....R: 60
  - b) { 2, 2, 2 }........ no excel : =(2*2*2)^(1/3) ....R: 2
  - c) { 1, 4, 16, 64 }........ no excel : =(1*4*16*64)^(1/4) ....R: 8

SEPARATRIZES (quartis, decis e percentis)

- Além das medidas de posição que estudamos (médias), há outras que, consideradas individualmente, não são medidas de tendência central, mas estão ligadas à **mediana** relativamente à sua característica de separar a série em duas partes que apresentam o mesmo número de valores.
- Essas medidas - os **quartis**, os **decis** e os **percentis** - são, juntamente com a **mediana**, conhecidas pelo nome genérico de separatrizes.

QUARTIS

- Denominamos **quartis** os valores de uma série que a dividem em quatro partes iguais. Precisamos portanto de 3 quartis (**Q1** , **Q2** e **Q3**) para dividir a série em quatro partes iguais.
  - Obs: O quartil 2 (Q2) sempre será igual a mediana da série.
- O método mais prático é utilizar o princípio do cáculo da mediana para os 3 quartis. Na realidade serão calculadas " **3 medianas** " em uma mesma série.

QUARTIS

Ex. 1: Calcule os quartis da série: { 5, 2, 6, 9, 10, 13, 15 }

- O primeiro passo a ser dado é o da ordenação (crescente ou decrescente) dos valores: { 2, 5, 6, 9, 10, 13, 15 }
- O valor que divide a série acima em duas partes iguais é igual a 9, logo a Md = 9 que será = Q2. Q1 Q3
- Temos agora {2, 5, 6 } e {10, 13, 15 } como sendo os dois grupos de valores iguais proporcionados pela mediana (quartil 2). Para o cálculo do quartil 1 e 3 basta calcular as medianas das partes iguais provenientes da verdadeira Mediana da série (quartil 2).

QUARTIS

Ex. 2: Calcule os quartis da série: { 1, 1, 2, 3, 5, $\boxed{5, 6,}$ 7, 9, 9, 10, 13 }

Quartil 2 = Md = (5+6)/2 = 5,5

- O quartil 1 será a mediana da série à esquerda de Md : { 1, 1, $\boxed{2, 3,}$ 5, 5 }

  Q1 = (2+3)/2 = 2,5

- O quartil 3 será a mediana da série à direita de Md : {6, 7, $\boxed{9, 9,}$ 10, 13 }

  Q3 = (9+9)/2 = 9

DECIS e PERCENTIS

DECIS

A definição dos decis obedece ao mesmo princípio dos quartis. Indicamos os decis : D1, D2, ... , D9. Deste modo precisamos de 9 decis para dividirmos uma série em 10 partes iguais.

PERCENTIL ou CENTIL

Denominamos percentis ou centis como sendo os noventa e nove valores que separam uma série em 100 partes iguais. Indicamos: P1, P2, ... , P99. É evidente que P50 = Md ; P25 = Q1 e P75 = Q3.

Desvio Padrão

É a medida de dispersão mais geralmente empregada, pois leva em consideração a totalidade dos valores da variável em estudo. O desvio padrão baseia-se nos desvios em torno da média aritmética e a sua fórmula básica pode ser traduzida como : a raiz quadrada da média aritmética dos quadrados dos desvios e é representada por σ .

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{n}}$$

x_i são os valores da variável

média

Desvio Padrão

Exemplo:

Calcular o desvio padrão da população representada por:

{ - 4 , -3 , -2 , 3 , 5}

Xi	$\overline{X}$	$(Xi - \overline{X})$	$(Xi - \overline{X})^2$
- 4	- 0,2	- 3,8	14,44
- 3	- 0,2	- 2,8	7,84
- 2	- 0,2	- 1,8	3,24
3	- 0,2	3,2	10,24
5	- 0,2	5,2	27,04
		E =	62,8

Sabemos que n = 5 e 62,8 / 5 = 12,56.

A raiz quadrada de 12,56 é o desvio padrão = **3,54**

Desvio Padrão

Obs: Quando nosso interesse não se restringe à descrição dos dados mas, partindo da amostra, visamos tirar inferências válidas para a respectiva população, convém efetuar uma modificação, que consiste em usar o divisor n - 1 em lugar de n. A fórmula ficará então:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}$$

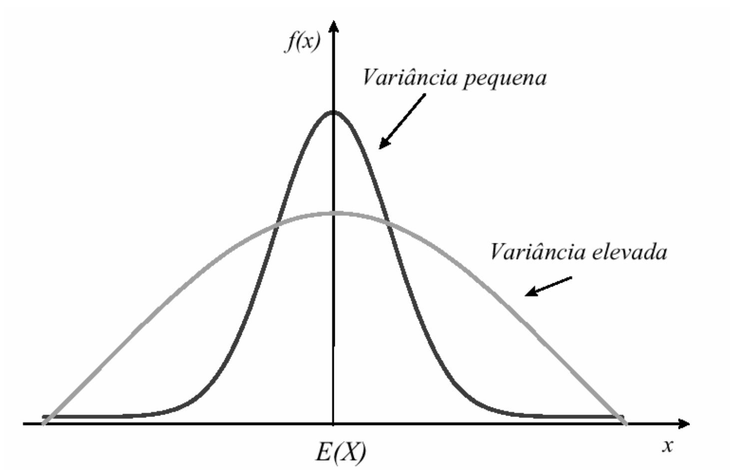
x_i são os valores da variável

média

VARIÂNCIA

É o desvio padrão elevado ao quadrado e é simbolizado por σ^2. A variância é uma medida que tem pouca utilidade como estatística descritiva, porém é extremamente importante na inferência estatística e em combinações de amostras.

VARIÂNCIA



Tamanho da amostra

- O tamanho teórico (n)de uma amostra para obter-se uma incerteza Δ na forma:

$$x = \bar{x} \pm \Delta$$

média

pode ser obtido por:

Depende do nível de confiança

Intervalo de confiança

$$\sqrt{n} = \frac{z\sigma}{\Delta}$$

Número de amostras ou observações

Desvio padrão das observações

Tamanho da amostra

z

Intervalo de confiança	Nível de confiança %
3,30	99,9
3,0	99,7
2,57	99,0
2,0	95,4
1,96	95,0
1,65	90,0
1,0	68,3

Tamanho da amostra

Exemplo

Suponha que um série de medidas é efetuada com um desvio padrão de ± 0,5 mm (devido a precisão do instrumento e variabilidade do experimento). Quantas medições são necessárias para estabelecer um valor médio uma incerteza (Δ) de 0,2 mm, na forma:

$$x = \bar{x} \pm 0{,}2mm$$

com um nível de confiança de 99,9 %.

$z = 3{,}30 \leftarrow$ Valor da tabela

$$\sqrt{n} = \frac{z\sigma}{\Delta} \longrightarrow \sqrt{n} = \frac{3{,}30.0{,}5}{0{,}2} \longrightarrow n = 68{,}05 \longrightarrow n = 69 amostras$$