

Média aritmética: Para dados isolados, a média aritmética indicada por \bar{x} . Sua fórmula é

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

Se os dados estiverem agrupados em classes, a fórmula a ser usada é

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

Aqui, x_i representa o ponto médio de cada classe ($superior + inferior \div 2$) e este valor deve ser multiplicado por n_i que é a quantidade de elementos naquela classe

Mediana: A mediana de um conjunto de valores é o valor do meio, depois de ordenados os valores segundo o critério a examinar. Para dados isolados, se o número de elementos considerados for ímpar, a obtenção do elemento central é imediata. Se o número for par a mediana deve ser calculada pela média aritmética dos dois elementos centrais. O símbolo da mediana é M_d . Para dados agrupados em classes, agora a fórmula é

$$M_d = l_{M_d} + \left(\frac{\frac{\sum n_i}{2} - n_{aM_d}}{n_{M_d}} \right) \cdot h_{M_d}$$

sendo

l_{M_d} limite inferior da classe mediana

n_{aM_d} n_i acumulado anterior ao l_{M_d}

n_{M_d} n_i da própria classe mediana

h_{M_d} tamanho da classe mediana

Moda: A moda, representada por M_o de um conjunto de dados é o valor que aparece mais vezes, ou seja aquele ao qual esteja associado a frequência absoluta mais alta.

Para dados não agrupados, não há fórmula, basta olhar os dados e ver qual o mais frequente.

Para dados agrupados em classes a fórmula, devida a Czuber, é

$$M_o = l_{M_o} + \left(\frac{n_{aM_o} - n_{M_o}}{n_{aM_o} + n_{pM_o} - 2n_{M_o}} \right) \cdot h_{M_o}$$

onde

l_{M_o} limite inferior da classe de maior frequência absoluta

n_{aM_o} n_i anterior ao maior n_i

n_{M_o} n_i maior n_i da tabela

n_{pM_o} n_i posterior ao maior n_i

h_{M_o} tamanho da classe modal

Medidas de Variabilidade

Descreve-se aqui algumas medidas que indicam as oscilações de uma variável, daí o nome. Como vimos acima a média aritmética indica onde as medidas tendem a se agrupar, mas a pergunta agora é quanto elas variam em torno àquele valor.

Desvio padrão: É uma quantidade que mede a amplitude da variação das medidas em torno da média aritmética. É indicado por $s(x)$, que se lê *s de x*. O desvio padrão indica qual o valor médio em que a medida varia em relação a média da população.

As fórmulas (isolado e agrupado) do d.p. são:

$$dp(x) = + \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \quad dp(x) = + \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot n_i}{\sum_{i=1}^n n_i}}$$

onde x_i vão ser os pontos médios de cada classe e as diferenças DEPOIS de elevadas ao quadrado é que deverão ser multiplicadas pelos respectivos n_i .

Variância: Indicado por $s^2(x)$, que se lê *s-quadrado de x*, é exatamente o quadrado do desvio padrão. Para os dados agrupados a variância segue a fórmula vista acima para o desvio padrão de dados agrupados, mas sem extrair a última raiz.

Para você fazer

Você deve ler o arquivo

FT51D01.myd

e nele obter os dados para todos os exercícios aqui propostos

1. Suponha um levantamento feito com 1000 alunos da Universidade Positivo. Para cada aluno, questionou-se a idade, sexo, escolaridade do pai e da mãe, renda familiar e distância da residência até o local de aula. Um exemplo do tipo de dado é

23 1 4 3 15000 14.4

A coluna de sexo deve ser interpretada como 1=masculino e 2=feminino e 3=indeterminado/desconhecido. As colunas de escolaridade como 1=sem escolaridade, 2=ensino fundamental, 3=ensino médio, 4= ensino superior, 5=pós graduado, 6=indeterminado ou desconhecido. A renda média é em reais por mês e a distância em quilômetros. A partir desse universo (1000 registros), achar:

- idade média das pessoas com sexo=2
- renda familiar média
- moda da escolaridade do pai
- desvio padrão da distância entre a casa e a escola

2. No ano passado, houve um levantamento de dados meteorológicos na nossa cidade. Registraram-se 365 ocorrências das seguintes informações: dia do mês do levantamento ($1 \leq dia \leq 31$)
 mês do levantamento ($1 \leq mes \leq 12$)
 horas de insolação: quantas horas fez sol nesse dia
 temperatura máxima: em graus Celsius nesse dia
 temperatura mínima: em graus Celsius nesse dia
 velocidade máxima do vento: em Km/h nesse dia
 precipitação: em milímetros nesse dia

Achar a amplitude de temperatura média no ano. Para tanto, calcular a amplitude diária e depois achar a média. Achar a mediana da velocidade do vento. Achar a variância da insolação diária.

3. Para este exercício os dados estão agrupados por faixa de salário. Trata-se de um levantamento de remuneração de salário. Os dados têm as seguintes informações:
 estado da federação: Duas letras indicativas do estado
 valor inicial da faixa de salário: em reais
 valor final da faixa de salário: em reais
 quantidade de assalariados nessa faixa: em unidades

Calcular:

Mediana dos salários no estado Paraná

Moda dos salários em Santa Catarina

Desvio Padrão nos salários em São Paulo

Exemplo Para testar seu programa rode-o com o arquivo `exet51.myd` e tente obter as respostas:

23.9 6107.1 2 2.8 6.4 12.0 15.4
 3867.9 1394.8 3229.1

Responda aqui :

1a. Idade média S=2	
1b. Renda familiar média	
1c. Moda esc. do pai	
1d. DP da distância escola-casa	
2a. ampl. temp. média ano	
2b. Mediana velocidade do vento	
2c. Variância da insolação	
3a. Mediana salários PR	
3b. Moda salários em SC	
3c. DP salários em SP	

Para as respostas não inteiras, use 1 casa decimal

