

| | | |
|---|--|---|
| FACULDADE DE TECNOLOGIA SENAC RIO | | |
| Curso: Análise e Desenvolvimento de Sistemas | | Semestre letivo: 2024.1 |
| Unidade Curricular: Estatística Aplicada | | Módulo: 3 |
| Professor: Agnaldo Cieslak | | Data: 16.05.2024 |
| Competências a serem avaliadas: <ul style="list-style-type: none"> Desenvolver sistemas computacionais aplicando boas práticas de Qualidade de Software | | Indicadores de Competência: Aplica as técnicas de tratamento estatístico de dados e informações importantes para o processo de tomada de decisão. |
| Aluno: Erick Calazães da Silva | | Conceito: |

Tarefa 6 – Workshop de exercícios revisórios

Orientação: os exercícios devem ser elaborados com o demonstrativo dos passos de sua resolução.

- 21) Numa escola é adotado o seguinte critério: a nota da primeira prova é multiplicada por 1, a nota da segunda prova é multiplicada por 2 e a nota da última prova é multiplicada por 3. Os resultados, após somados, são divididos por 6. Se a média obtida por esse critério for maior ou igual a 6,5, o aluno é dispensado das atividades de recuperação. Suponha que um aluno tenha tirado 6,3 na primeira prova e 4,5 na segunda. Quanto precisará tirar na terceira para ser dispensado da recuperação?

R: $3x = (6,5 \cdot 6) - (6,3 \cdot 1) - (4,5 \cdot 2) / x = 23,7 / 3 / x = 7,9$

- Quais são os tipos de amostragem utilizados em cada situação apresentada abaixo:
 - Ao escalar uma comissão para atuar em determinado projeto, uma empresa decidiu selecionar aleatoriamente 4 pessoas brancas, 3 pardas e 4 negras.

R: Amostragem Estratificada Proporcional

b) Uma professora escreve o nome de todos os seus alunos em pedaços de papel e coloca em uma caixa. Depois de misturá-los, sorteia 10 nomes.

R: Amostragem Aleatória Simples

c) Um administrador de uma sala de cinema faz uma pesquisa com as pessoas que estão na fila de espera para comprar ingresso, entrevistando uma pessoa a cada 10 presentes na fila.

R: Amostragem por conveniência

d) Deseja-se selecionar uma amostra de domicílios da cidade de São Paulo. As ruas estão identificadas pelas letras de A a F. As casas de cada rua estão identificadas pelo nome da rua, seguido por um número. Primeiro foram sorteadas duas ruas (B e F) e depois, foram selecionados ao acaso 50% dos domicílios de cada rua.

R: Amostragem por Conglomerado

- Imagine que você tem 500 cadastros arquivados em sua empresa e você quer uma amostra de 2% desses cadastros. Como você obterá uma amostra sistemática?

R: Ordenando os cadastros de forma aleatória e selecionando 1 elemento aleatoriamente a partir da tabela de números aleatórios a cada 50 ($k=N/n$; $X(TNA)$, $X+K$, $X+2K$, ...) até que sejam selecionados 10 elementos ($2\% \cdot 500$).

4. Uma federação encomendou uma pesquisa na cidade de “Limas” sobre a formação escolar e técnica dos trabalhadores da indústria. A cidade tem cerca de 25000 pessoas trabalhando em 160 indústrias. As indústrias foram classificadas de acordo com o número de trabalhadores e mostrados na tabela abaixo:

| Nº Trabalhadores | Porte da indústria | Nº de indústrias |
|------------------|--------------------|------------------|
| 1 a 50 | pequena | 100 |
| 51 a 400 | média | 40 |
| Acima de 400 | grande | 20 |

As indústrias foram cadastradas de 001 a 160. O pesquisador deve entrevistar todos os trabalhadores de 5 indústrias.

Qual a característica desta população?

Qual técnica de amostragem deve ser utilizada?

Descreva sucintamente o passo a passo para obter a amostragem correta, apresentando o seu resultado. (quais empresas deveriam ser selecionadas)?

R: A característica da população a ser estudado (parâmetro) é a formação escolar e técnica desses trabalhadores. A técnica a ser utilizada é a Amostragem Estratificada Proporcional, de modo que possa se garantir uma representação adequada da população. Selecionaria-se então indústrias aleatórias seguindo a amostragem abaixo:

| | | | | | | |
|---|-----|------|-------|----|-------|-----------------------|
| P | 100 | /160 | 0,625 | *5 | 3,125 | 3 indústrias P |
| M | 40 | /160 | 0,25 | *5 | 1,25 | 1 indústria M |
| G | 20 | /160 | 0,125 | *5 | 0,625 | 1 indústria G |

Total 160

Amostra 5

9. Uma faculdade apresentava, no final do ano, o seguinte quadro:

| PERÍODOS | MATRÍCULAS | |
|--------------|--------------|--------------|
| | MARÇO | NOVEMBRO |
| 1º | 480 | 475 |
| 2º | 458 | 456 |
| 3º | 436 | 430 |
| 4º | 420 | 420 |
| Total | 1.794 | 1.781 |

- a. Calcule a taxa de evasão por período.
b. Calcule a taxa de evasão da faculdade.

| | Matrículas | | | | |
|----------|------------|----------|---------------------------------|------|---------------------|
| Períodos | Março | Novembro | | | |
| 1º | 480 | 475 | TxEvasão 1 = (480 -475) / 480 = | 0,01 | *100% = 1,04 |
| 2º | 458 | 456 | TxEvasão 2 = (458 -456) / 458 = | 0,00 | *100% = 0,44 |
| 3º | 436 | 430 | TxEvasão 3 = (436 -430) / 436 = | 0,01 | *100% = 1,38 |
| 4º | 420 | 420 | TxEvasão 4 = (420 -420) / 420 = | 0,00 | *100% = 0 |
| Total | 1794 | 1781 | Txde Evasão Média = SomaTx / 4 | | = 0,71 |

10) Medidas as estaturas de 1.017 indivíduos, obtivemos $X = 162,2$ cm e $s = 8,01$ cm. O peso médio desses mesmos indivíduos é 52 kg, com um desvio padrão de 2,3 kg. Esses indivíduos apresentam maior variabilidade em estatura ou em peso?

| $\%CV = (S / X) * 100$ | S | X | %CV |
|------------------------|------|-------|-------------|
| Estatura | 8,01 | 162,2 | 4,94 |
| Peso | 2,3 | 52 | 4,42 |

R: A estatura apresentou maior variabilidade que o peso.

17) Numa competição de salto triplo, três atletas disputavam apenas uma vaga para uma olimpíada entre faculdades de uma cidade. Para ser aprovado deveria descartar o pior resultado e obter a melhor média. Cada atleta fez 4 tentativas obtendo os seguintes resultados:

| | | | | |
|-------------------|---------|---------|---------|---------|
| Atleta I | 16,50 m | 15,81 m | 16,42 m | 16,12 m |
| Atleta II | 13,90 m | 17,01 m | 16,82 m | 15,10 m |
| Atleta III | 15,70 m | 16,02 m | 16,95 m | 17,00 m |

a) Qual deles foi aprovado? R: O atleta III.

b) Qual deles foi o mais regular nessas quatro tentativas? R: O atleta I.

c) Considerando todas as medidas, calcule a média e desvio padrão de cada atleta e em conjunto de todos os atletas e compare através do cálculo do coeficiente de variação, comentando os resultados. R: Comparando os dados obtidos, entende-se que o atleta I foi o bastante consistente em seu desempenho, seguido do atleta III e do atleta II, respectivamente.

| Atleta | Tentativa 1 | Tentativa 2 | Tentativa 3 | Tentativa 4 | Média descartando a menor | Média verdadeira |
|--------|-------------|-------------|-------------|-------------|---------------------------|------------------|
| 1 | 16,5 | 15,81 | 16,42 | 16,12 | 16,35 | 16,21 |
| 2 | 13,9 | 17,01 | 16,82 | 15,1 | 16,31 | 15,71 |
| 3 | 15,7 | 16,02 | 16,95 | 17 | 16,66 | 16,42 |

| | | | | | | | |
|----------------|------|------|------|------|-------|--------------|------|
| $\sigma_1^2 =$ | 0,08 | 0,16 | 0,04 | 0,01 | 0,074 | $\sigma_1 =$ | 0,27 |
| $\sigma_2^2 =$ | 3,27 | 1,70 | 1,24 | 0,37 | 1,643 | $\sigma_2 =$ | 1,28 |
| $\sigma_3^2 =$ | 0,51 | 0,16 | 0,28 | 0,34 | 0,324 | $\sigma_3 =$ | 0,57 |

| | |
|-----------------------------------|------|
| $\%CV_1 = (\sigma_1 / u) * 100 =$ | 1,68 |
| $\%CV_2 = (\sigma_2 / u) * 100 =$ | 8,16 |
| $\%CV_3 = (\sigma_3 / u) * 100 =$ | 3,47 |

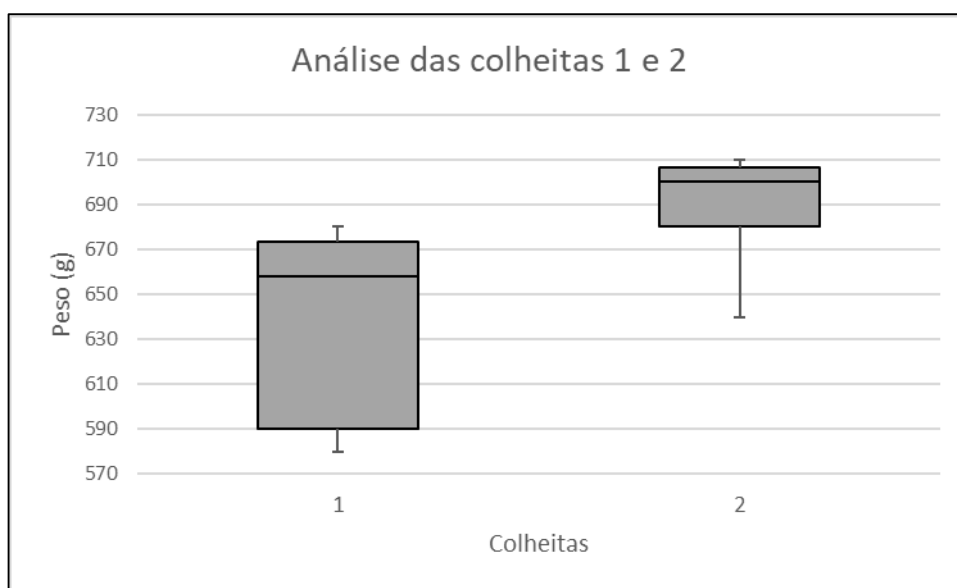
8. Num laboratório de pesquisa foram medidos os pesos de 16 frutos que fazem parte de um estudo para aumento de produção. Os resultados em gramas foram apresentados na tabela abaixo:

| Colheita | Frutos – peso em g | | | | | | | |
|----------|--------------------|-----|-----|-----|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Primeira | 648 | 595 | 668 | 580 | 672 | 585 | 675 | 680 |
| Segunda | 680 | 640 | 700 | 710 | 708 | 680 | 705 | 700 |

Determine, com uma casa decimal: a) Média b) Mediana c) Moda d) Separtriz (Q1, Q3, D6) e) Amplitude f) Variância g) Desvio-padrão h) Coeficiente de variação i) Faça a representação gráfica das medidas estatísticas (Box-Plot) j) Analise os resultados e registre suas conclusões.

| Colheita | Média | Mediana | Moda | Q1 | Q3 | D6 |
|----------|-------|---------|----------|-----|-------|-----|
| Primeira | 637,9 | 658 | n/a | 590 | 673,5 | 668 |
| Segunda | 690,4 | 700 | 680; 700 | 680 | 706,5 | 700 |

| Colheita | Amplitude | Variância | Desvio padrão | Coeficiente de variação |
|----------|-----------|-----------|---------------|-------------------------|
| Primeira | 100 | 1902 | 43,6 | 6,8 |
| Segunda | 70 | 550 | 23,4 | 3,4 |



R: Analisando o gráfico podemos concluir que a colheita 2 foi muito mais consistente em termos de peso dos frutos analisados. Podemos concluir também que os frutos colhidos na colheita 2 foram mais pesados que os da colheita 1, o que provavelmente no caso demonstraria uma maior qualidade da colheita e da produção no geral.

Amplitude $R = X_{(n)} - X_{(1)}$

Variância Amostral $s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}$

Desvio Padrão amostral $s = \sqrt{s^2} = \sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}}$

Coeficiente de variação $\%CV = \frac{S}{\bar{X}} \cdot 100$

Indicador de homogeneidade de dados
CV \leq 15% \rightarrow baixa dispersão em relação à média
15% < CV < 30% \rightarrow média dispersão em relação à média
CV \geq 30% \rightarrow alta dispersão em relação à média