Aula 4 - Medidas resumo

PhD. Wagner Hugo Bonat

Laboratório de Estatística e Geoinformação-LEG
Universidade Federal do Paraná

1/2017

Medidas de posição para um conjunto de dados

- Média:
$$\bar{x}_{obs} = \frac{x_1 + x_2 \ldots x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}.$$
-
$$\bar{x}_{obs} = \frac{n_1 x_1 + n_2 x_2 + \ldots n_k x_k}{n_1 + n_2 + \ldots n_k} = \frac{\sum_{i=1}^{k} n_i x_i}{n}.$$
- Mediana (md_{obs}): Valor que ocupa a posição central dos dados ordenados.
- Moda (mo_{obs}): valor mais frequente.

Laboratório de Estatística e Geoinformação

Exemplo 4.1

- Suponha que parafusos a serem utilizados em tomadas elétricas são embalados em caixas rotuladas como contendo 100 unidades. Em uma construção, 10 caixas de um lote tiveram o número de parafusos contados, fornecendo os valores 98, 102, 100, 100, 99, 96, 95, 99 e 100. Calcular média, mediana e moda.
- $\bar{x}_{obs} = 98.6$.
- $md_{obs} = 99$.
- $mo_{obs} = 100$.

Laboratório de Estatística e Geoinformação

Exemplo 4.2

- Nas caixas de parafusos do Examplo 4.1, vamos admitir um custo de c por parafuso e de e pela embalagem da caixa. Desejamos calcular as medidas de posição do custo total (T), definido como a soma dos custos dos parafusos e da embalagem.

Exemplo 4.3

- Foram coletadas 150 observações da variável X, representando o número de vestibulates FUVEST (um por ano) que um mesmo estudande prestou. Assim, foi observado que 75 estudantes prestaram vestibular FUVEST, uma única vez, e assim por diante. Os dados estão na tabela abaixo:

X	n_i
1	75
2	47
3	21
4	7

- Suponha ainda que o interesse é estudar o gasto dos alunos associado com as despesas do vestibular. Para simplificar, suponha que se atribui para cada aluno, uma despesa fixa de 1300,00 relativa a preparação e mais 50 para cada vestibular prestado. Calcule as medidas de posição central.

Laboratório de Estatística e Geoinformação

Exemplo 4.4

- Um estudante está procurando um estágio para o próximo ano. As companhias A e B têm programas de estágios e oferecem uma remuneração por 20 horas semanais com as seguintes características.

Companhia	A	B
média	2.5	2.0
mediana	1.7	1.9
moda	1.5	1.9

- Qual companhia você escolheria?

Medidas de posição para va

- Valor esperado

$$E(X) = \sum_{i=1}^{k} x_o p_i.$$

- Mediana

$$P(X \geq Md) \geq 1/2 \quad \text{e} \quad P(X \leq Md) \leq 1/2.$$

- Moda

$$P(X = mo) = \max(p_1, p_2, \ldots, p_k).$$

Laboratório de Estatística e Geoinformação

Exemplo 4.5

- Conside a va X com a seguinte função discreta de probabilidade:

X	-5	10	15	20
p_i	0.3	0.2	0.4	0.1

Laboratório de Estatística e Geoinformação

Exemplo 4.6

- Conside uma va X com função de probabilidade dada por

X	2	5	8	15	20
p_i	0.1	0.3	0.2	0.2	0.2

- Calcule as medidas de posição para a va $Y = 5X - 10$.
- Veja resumo na Tabela 4.1 pg. 100.

Amplitude de uma variável em um conjunto de dados

- Amplitude: diferença entre o maior e menor valor do conjunto de dados.
- Exemplo 4.7: Numa classe com 12 alunos de um curso de inglês, os alunos indicaram o número de outras línguas que tinham alguma familiaridade. Os resultados ordenados foram: 0,0,0,0,1,1,1,1,1,2,2 e 4. Calcule as medidas de posição central e avalie algumas medidas de dispersão.
- Desvio mediano: $\frac{1}{n}\sum_{i=1}^{n}|x_i - md_{obs}|$.
- Desvio médio: $\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}_{obs}|$.

Laboratório de Estatística e Geoinformação

Variância e desvio-padrão em um conjunto de dados

- Variância $var_{obs} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_obs)^2$.
- Desvio-padrão $dp_{obs} = var_{obs}$.
- Exemplo 4.8: No Exemplo 4.1, foram fornecidas as quantidades de parafusos em 10 caixas de um lote. Calcule a variância e o desvio-padrão.

```
> obs <- c(98,102,100,100,99,97,96,95,99,100)
> n <- length(obs)
> media <- sum(obs)/n
> desvio <- obs - media
> desvio

 [1] -0.6  3.4  1.4  1.4  0.4 -1.6 -2.6 -3.6  0.4  1.4

> sum(desvio^2)

[1] 40.4

> sum(desvio^2)/n

[1] 4.04

> sum(obs^2)/n - media^2

[1] 4.04

> sqrt(sum(desvio^2)/n)

[1] 2.009975
```

Laboratório de Estatística e Geoinformação

Exemplo 4.9

- No Exemplo 4.3, definimos a quantidade D, despesa no vestibular obtida a partir de X pela expressão $D = 50X + 1300$ com X indicando o número de vestibulares prestados. Calcule a variância de D.
- Exemplo 4.10 tarefa de casa.

Variância de uma va

- $Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 p_i$.
- $Var(X) = E[(X - \mu)^2] = E(X^2) - E^2(X)$.
- Ver Tabelas resumo 4.2 e 4.3 pg. 111.

Exemplo 4.11

- Uma pequena cirurgia dentária pode ser realizada por três métodos diferentes cujos tempos de recuperação (em dias) são modelados pelas variáveis X_1, X_2 e X_3. Admita suas funções de probabilidades são dadas por

X_1	0	4	5	6	10
p_i	0.2	0.2	0.2	0.2	0.2

X_2	1	5	9
p_i	$1/3$	$1/3$	$1/3$

X_3	4	5	6
p_i	0.3	0.4	0.3

- Calcule as medidas de posição central e dispersão para cada va e decida sobre o método mais eficiente.

Laboratório de Estatística e Geoinformação

Exemplos

- Exemplo 4.14: Seja X com distribuição Bernoulli de parâmetro p. Calcule a esperança e a variancia de X.
- Exemplo 4.15: Seja X com distribuição Binomial parâmetros n e p. Calcule a esperança e a variancia de X.
- Ver resultados da Tabela 4.4 pg. 113.

Laboratório de Estatística e Geoinformação

Exercícios recomendados

- Seção 4.2 - 1,2,3,4 e 6.
- Seção 4.3 - 1,2,3,4,5 e 6.