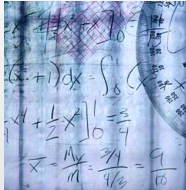


Estatística Aplicada

Noções Introdutórias e utilização do computador



Segunda aula - Estatística Descritiva I

Divisão da Estatística

- Estatística Descritiva
- Estatística Analítica (Inferência estatística/Estatística indutiva)

Estatística Descritiva

Pedro Menezes - 2008 2

Divisão da Estatística

Estatística Descritiva

Método quantitativo utilizado apenas para descrever os dados, mostrando seus subtipos, sua distribuição, frequência, média etc.

Estatística Descritiva

Pedro Menezes - 2008 3

Divisão da Estatística

Estatística Descritiva

Parâmetros médio de erro		
	Grupo controle	Grupo teste
Plano horizontal	45,8	48,7
Plano sagital mediano	52,3	46,5
0,5 kHz	43,3	48,01
2 kHz	46,4	49,5
4 kHz	46,9	71,9

Média das amplitudes das EOA-PD, por sexo

Pedro Menezes - 2008 4

Divisão da Estatística

Estatística Analítica

Método quantitativo utilizado comparar grupos e fazer generalizações, a partir de resultados obtidos.

Pedro Menezes - 2008 5

Divisão da Estatística

Estatística Analítica

Frequência	Valores de p
0,5kHz	0,621
1kHz	0,146
2kHz	0,197
4kHz	0,236
8kHz	0,216
16kHz	0,621

Salas	Ruído Médio
1º período	75,7 dB(A)
2º período	75,1 dB(A)
3º período	75,7 dB(A)

Teste t de Student: entre 1 e 2, $p=0,657$, entre 1 e 3, $p=0,832$ e entre 2 e 3, $p=0,220$

Resultado do teste t de Student, entre os participantes, com e sem venda, por frequência

Pedro Menezes - 2008 6

Tipos de Dados

1. Dados Categóricos

- Dados Nominais
- Dados Ordinais

2. Dados Numéricos

- Discretos
- Contínuos

Estatística Descritiva

Pedro Menezes - 2008 7

Tipos de Dados

Dados categóricos: Nominais

Dados amorfos, que não estão estruturados numericamente, sem ordem. Não são mensurados, apenas contados.

Exemplos: sexo, raça, cor, tipo de parto, etc.

Podem ser codificados por números, por exemplo,

- Masculino = 1
- Feminino = 2.

Estatística Descritiva

Pedro Menezes - 2008 8

Tipos de Dados

Dados Nominais

Forma comum de apresentação:

- Diagramas de barras
- Pizzas
- Tabelas.

Estatística Descritiva

Pedro Menezes - 2008 9

Tipos de Dados

Dados categóricos: Ordinais

Dados que se distribuem por categorias que têm uma ordem.

Exemplos:

- salinidade: baixa, média, alta
- abundância: dominante, abundante, freqüente, ocasional, raro
- intensidade: forte, média, fraca

Estadística Descritiva

Pedro Menezes - 2008 10

Tipos de Dados

Dados Ordinais

Forma comum de apresentação:

- Mediana
- Proporções
- Percentagens

Estadística Descritiva

Pedro Menezes - 2008 11

Tipos de Dados

Dados numéricos: Discretos

Dados que são contagens representadas por números inteiros.

Exemplos:

- número de ovos em um galinheiro
- número de episódios de dor de ouvido
- número de casos de determinada patologia

Estadística Descritiva

Pedro Menezes - 2008 12

Tipos de Dados

Dados numéricos: Contínuos

Dados com números intrinsecamente significantes, onde as diferenças possuem sempre o mesmo significado. Medidas numa escala contínua.

Exemplos:

- Peso, volume, área, massa, tamanho da cabeça, etc.

Pedro Menezes - 2008 13

Tipos de Dados

Rigidez

As distinções são menos rígidas do que parecem.

- Idade = variável contínua, mas se a idade for registrada pelo ano mais próximo, podemos tratá-la como discreta.
- Idade também pode ser categórica, se separarmos a amostra em "crianças", "adultos jovens", "idade média", "velhos".

No entanto, em geral é recomendado manter os dados em sua forma original, categorizando os dados somente para propósitos de apresentação.

Pedro Menezes - 2008 14

Processos de medida

Índices de exatidão

PRECISÃO: Uma medida é tão mais precisa quanto mais próxima estiver do valor médio da grandeza associada.

EXATIDÃO: Uma medida é tão mais exata quanto menor for o "vício" da medida, ou seja, a diferença entre o valor mais provável (valor médio) encontrado e o verdadeiro valor da grandeza medida, suposto teoricamente conhecido.

Pedro Menezes - 2008 15

Processos de medida

Índices de exatidão

Atirador 1

$\bar{x} = 2$

Atirador 2

$\bar{x} = 1,5$

Quem foi mais preciso: o atirador 1 ou o 2?
 Quem foi mais exato: o atirador 1 ou o 2?

Pedro Menezes - 2008 16

População e amostra

População

Coleção de unidades individuais, que podem ser pessoas ou resultados experimentais, com uma ou mais características comuns, que se pretendem estudar.

Pedro Menezes - 2008 17

População e amostra

População


Exemplo 1: Relativamente à população constituída pelos alunos do 1º ano de um curso qualquer da uncial, podemos estar interessados em estudar as seguintes características populacionais:

- Altura
- Notas
- Avaliação auditiva

Pedro Menezes - 2008 18

População e amostra

Nem sempre é possível estudar exaustivamente todos os elementos da população!




1. A população pode ter dimensão infinita
Exemplo: População constituída pelas pressões atmosféricas, nos diferentes pontos de uma cidade.
2. O estudo da população pode levar à destruição da população
Exemplo: Fósforos que acedem em uma caixa de fósforos.

Pedro Menezes - 2008 19

População e amostra

Nem sempre é possível estudar exaustivamente todos os elementos da população!



3. O estudo da população pode ser muito dispendioso
Exemplo: Sondagens exaustivas de todos os eleitores, sobre determinado candidato.

Pedro Menezes - 2008 20

População e amostra

Amostra

145	148
156	160
163	157
152	149
149	147

Conjunto de dados ou observações, recolhidos a partir de um subconjunto da população, que se estuda com o objetivo de tirar conclusões para a população de onde foi recolhida.

Pedro Menezes - 2008 21

Estatística Descritiva

População e amostra

Amostra

Exemplo 1: Relativamente à população das alturas dos alunos do 1º ano de um dos cursos da uncial, consideremos a seguinte **amostra**, constituída pelas alturas (em cm) de 20 alunos escolhidos ao acaso:

145, 163, 157, 152, 156, 149, 160, 157, 148, 147, 151, 152, 150, 148, 156, 160, 148, 157, 153, 162

Pedro Menezes - 2008 22

Estatística Descritiva

Análise descritiva dos dados

Resumindo numericamente

Para resumir numericamente dados quantitativos o objetivo é escolher medidas apropriadas de **locação** ("qual o tamanho dos números envolvidos?") e de **dispersão** ("quanta variação existe?") para os tipos de dados.

Pedro Menezes - 2008 23

Estatística Descritiva

Análise descritiva dos dados

- Medidas de tendência central
- Medidas de dispersão
- Separatrizes

Pedro Menezes - 2008 24

Estatística Descritiva

Análise descritiva dos dados

1. Medidas de tendência central

Vamos estudar seguidamente algumas medidas de localização, nomeadamente as que localizam o centro da amostra:

- Média
- Moda
- Mediana

Pedro Menezes - 2008 26

Estatística Descritiva

Análise descritiva dos dados

1. Medidas de tendência central: Média

A média amostral ou simplesmente média, que se representa por \bar{x} é uma medida de localização do centro da amostra, e obtém-se a partir da seguinte expressão:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n},$$

onde x_1, x_2, \dots, x_n representam os elementos da amostra e n a sua dimensão.

Pedro Menezes - 2008 26

Estatística Descritiva

Análise descritiva dos dados

1. Medidas de tendência central: Média

Ao determinar a média dos seguintes dados:
12.4 13.5 13.6 11.2 15.1 10.6 12.4 14.3 113.5

obteve-se o valor $\bar{x} = 24.1$

Pedro Menezes - 2008 27

Análise descritiva dos dados

1. Medidas de tendência central: Média

Cuidados com a média

12.4 13.5 13.6 11.2 15.1 10.6 12.4 14.3 513.5

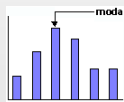
obteve-se o valor $\bar{x} = 68.5$

Pedro Menezes - 2008 28

Análise descritiva dos dados

1. Medidas de tendência central: Moda

Define-se como o valor que surge com mais frequência se os dados são discretos, ou, o intervalo de classe com maior frequência se os dados são contínuos.



Pedro Menezes - 2008 29

Análise descritiva dos dados

1. Medidas de tendência central: Moda

Esta medida é especialmente útil para reduzir a informação de um conjunto de dados qualitativos, apresentados sob a forma de nomes ou categorias, para os quais não se pode calcular a média e por vezes a mediana (se não forem susceptíveis de ordenação).

Pedro Menezes - 2008 30

Análise descritiva dos dados

1. Medidas de tendência central: Mediana

A mediana, m , é uma medida de localização do centro da distribuição dos dados, definida do seguinte modo:

Ordenados os elementos da amostra, a mediana é o valor que a divide ao meio, isto é, 50% dos elementos da amostra são menores ou iguais à mediana e os outros 50% são maiores ou iguais à mediana

Pedro Menezes - 2008 31

Análise descritiva dos dados

1. Medidas de tendência central: Mediana

Se n é ímpar, a mediana é o elemento médio.
Se n é par, a mediana é a semi-soma dos dois elementos médios.

Se se representarem os elementos da amostra ordenada com a seguinte notação:

$X_1:n, X_2:n, \dots, X_n:n$ então uma expressão para o cálculo da mediana será:

$$m = \begin{cases} X_{\frac{n+1}{2}:n} & \text{se } n \text{ é ímpar} \\ \frac{1}{2}(X_{\frac{n}{2}:n} + X_{\frac{n}{2}+1:n}) & \text{se } n \text{ é par} \end{cases}$$

Pedro Menezes - 2008 32

Análise descritiva dos dados

1. Medidas de tendência central: Mediana

$$m = \begin{cases} X_{\frac{n+1}{2}:n} & \text{se } n \text{ é ímpar} \\ \frac{1}{2}(X_{\frac{n}{2}:n} + X_{\frac{n}{2}+1:n}) & \text{se } n \text{ é par} \end{cases}$$

Exemplo: 5, 3, 3, 4, 2 (ímpar)

Ordenados: 2, 3, 3, 4, 5

$$m = X_{(n+1)/2} = (5+1)/2 = X_3$$

Pedro Menezes - 2008 33

Análise descritiva dos dados

1. Medidas de tendência central: Mediana

$$m = \begin{cases} x_{\frac{n+1}{2}} & \text{se } n \text{ é ímpar} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{se } n \text{ é par} \end{cases}$$

Média dos dois valores centrais = 3

Exemplo: 5, 1, 1, 4, 2, 5 (par)

Ordenados: 1, 1, 2, 4, 5, 5

$$m = \frac{1}{2}(x_{(n/2)} + x_{(n/2)+1}) = \frac{1}{2}(x_{(6/2)} + x_{(6/2)+1})$$

$$m = \frac{1}{2}(x_3) = x_{3,5}$$

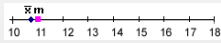
Pedro Menezes - 2008 34

Análise descritiva dos dados

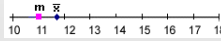
Média ou Mediana?

Um aluno do 1º ano obteve os seguintes dados:
10, 10, 10, 11, 11, 11, 11, 12

$\bar{x} = 10,75$ $m = 11$




Admitamos que uma das notas de 10 foi substituída por uma de 18. Neste caso a mediana continuaria a ser igual a 11, enquanto que a média subiria para 11.75!



Pedro Menezes - 2008 35

Análise descritiva dos dados

Média ou Mediana?



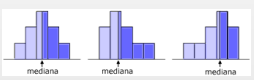
a mediana do peso destes atletas é 75kg, mas o peso médio é de 105kg!

Pedro Menezes - 2008 36

Análise descritiva dos dados

Média ou Mediana?

Dado um histograma é fácil obter a posição da mediana, pois esta está na posição em que passando uma linha vertical por esse ponto o histograma fica dividido em duas partes com áreas iguais.



Como medida de localização, a mediana é mais resistente do que a média, pois não é tão sensível aos dados.

Pedro Menezes - 2008 37

Análise descritiva dos dados

Média ou Mediana?

- Quando a distribuição é simétrica, a média e a mediana coincidem.
- A mediana não é tão sensível, como a média, às observações que são muito maiores ou muito menores do que as restantes (outliers). Por outro lado a média reflete o valor de todas as observações.

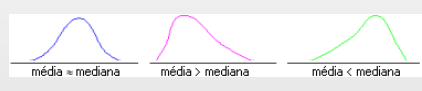
Assim, não se pode dizer em termos absolutos qual destas medidas de localização é preferível, dependendo do contexto em que estão a ser utilizadas.

Pedro Menezes - 2008 38

Análise descritiva dos dados

Média ou Mediana?

Resumindo, como a média é influenciada quer por valores muito grandes, quer por valores muito pequenos, se a **distribuição dos dados**:



Pedro Menezes - 2008 39

Análise descritiva dos dados

Exercício 1:

Calcular média, moda e mediana deste conjunto de dados:

1, 8, 4, 2, 2, 3, 1, 7, 3, 5, 3

Lembrete:
Moda: valor mais freqüente
Mediana: o valor central dos dados ordenados

Estadística Descritiva

Pedro Menezes - 2008 40

Análise descritiva dos dados

Resposta 1

Calcular média, moda e mediana deste conjunto de dados:

1, 1, 2, 2, 3, 3, 3, 4, 5, 7, 8

Média: 3,5
Moda: 3
Mediana: 3

Estadística Descritiva

Pedro Menezes - 2008 41

Análise descritiva dos dados

Exercício 2:

Calcular média, moda e mediana deste conjunto de dados:

1, 4, 2, 2, 3, 1, 7, 8, 5, 20

Estadística Descritiva

Pedro Menezes - 2008 42

Análise descritiva dos dados

Resposta 2

Calcular média, moda e mediana deste conjunto de dados:

1, 1, 2, 2, 3, 4, 5, 7, 8, 20

Média: 5,3
 Moda: 1 e 2
 Mediana: 3,5

Pedro Menezes - 2008 43

Análise descritiva dos dados

Exercício 3:

Suponha que os seguintes dados representam a concentração de uma enzima no sangue (mg/ml) da população humana em uma cidade A. Calcular média, moda e mediana, da amostra A, usando a opção Estatística Descritiva do Excel.

A =

8	8	6	6	6	7	8	0	1	1	4	3	5	5	7	9	9	2	2	2
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Pedro Menezes - 2008 44
