Aula de Exercícios

Organização: Rafael Tovar *Digitação*: Guilherme Ludwig

Mediana e Moda

Mediana: Dado que está na posição central, ou média dos valores das posições centrais, dependendo de n (o número de observações) ser par ou ímpar.

Moda: Valor mais frequente. Se os dados estão tabelados, pode-se aproximar pelo x^* (o ponto médio) do intervalo com o maior f_j. Porém deve-se ter em conta que os dados assim apresentados podem não ter uma moda, e logo esta é uma aproximação imperfeita.

Dados Tabelados

Considere os dados, de *Morettin & Bussab*, Estatística Básica.

N^o	Estado Civil	Instrução	Filhos	Salário	Idade	Região
1	solteiro	1^o grau	-	4,00	23 03	interior
2	casado	1^o grau	1	4,56	32 10	capital
3	casado	1^o grau	2	5,25	36 05	capital
4	solteiro	2^o grau	-	5,73	20 10	outro
5	solteiro	1^o grau	-	6,26	40 07	outro
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
36	casado	superior	3	23,30	42 02	interior

Dados Tabelados

Em forma de tabela, o salário pode ser apresentado como:

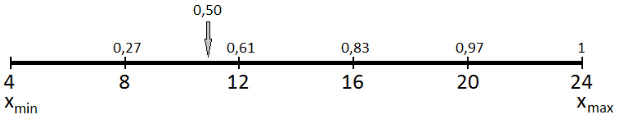
Intervalo	x^*	n	f	N	F
$(4 - 8]$	6	10	10	0,27	0,27
$(8 - 12]$	10	12	22	0,33	0,61
$(12 - 16]$	14	8	30	0,22	0,83
$(16 - 20]$	18	5	35	0,14	0,97
$(20 - 24]$	22	1	36	0,03	1

Dados Tabelados - Medidas de Posição Central e Dispersão

Dados na Lista	Dados na Tabela
$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$	$\bar{x} = \sum_{j=1}^{k} x_j^* \frac{n_j}{n} = \sum_{j=1}^{k} x_j^* f_j$
$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$	$s^2 = \sum_{j=1}^{k} (x_j^* - \bar{x})^2 \frac{n_j}{n}$
$s = \sqrt{s^2}$	$s = \sqrt{s^2}$
$CV = \frac{s}{\bar{x}}$	$CV = \frac{s}{\bar{x}}$

onde k é o número de intervalos de classe, e x_j^* é o ponto médio do intervalo de classe $(a_j - b_j]$.

Dados Tabelados - Cálculo da Mediana



Podemos obter o valor da mediana, ν, com uma regra de três:

$$\frac{12-8}{0{,}61-0{,}27} = \frac{\nu - 8}{0{,}5-0{,}27} \Leftrightarrow \nu = 10{,}7059$$

Distribuições Bivariadas

		Y				
		1	2	$\dots$	r	
	1	a_{11}	a_{12}	$\cdots$	a_{1r}	$\sum_{j=1}^{r} a_{2j}$
X	2	a_{21}	a_{22}	$\cdots$	a_{2r}	$\sum_{j=1}^{r} a_{2j}$
	$\vdots$	$\vdots$		$\ddots$	$\vdots$	$\vdots$
	k	a_{k1}	a_{k2}	$\cdots$	a_{kr}	$\sum_{j=1}^{r} a_{kj}$
		$\sum_{i=1}^{k} a_{i1}$	$\sum_{i=1}^{k} a_{i2}$	$\cdots$	$\sum_{i=1}^{k} a_{ir}$	$\sum_{i=1}^{k} \sum_{j=1}^{r} a_{ij}$

Distribuições Bivariadas

Exercício:

- Qual a proporção de indivíduos que tem categoria 1 de Y e categoria 2 de X?
- Qual a proporção de indivíduos que tem categoria 2 de Y, entre o total?
- Qual a proporção de indivíduos que tem categoria 2 de X, entre o total?
- Entre os elementos que tem a categoria r de Y, que proporção tem a categoria k de X?

Distribuições Bivariadas - Exemplo I

Considere os intervalos de classe para salário e a ditribuição covariada com a variável procedência. A tabela a seguir mostra a frequencia de cada classe:

	Capital	Interior	Outro	Marginal
$(4-8]$	4	3	3	10
$(8-12]$	3	4	6	13
$(12-16]$	1	3	3	7
$(16-20]$	3	1	1	5
$(20-24]$	0	1	0	1
Marginal	11	12	13	36

Distribuições Bivariadas - Exemplo II

Considere um estudo com 19476 homens com 50 anos ou mais, que foram submetidos a dois testes de rastreamento de câncer de próstata.

- *PSA:* Antígeno prostático específico.
- *DRE:* Toque retal (Digital Rectal Examination).
- *Confirmação:* biópsia.

Distribuições Bivariadas - Exemplo II

Pacientes com câncer				Pacientes sem câncer			
	DRE+	DRE-			DRE+	DRE-	
PSA+	189	292	481	PSA+	141	755	896
PSA-	145	1255	1400	PSA-	1002	15697	16699
	344	1547	1881		1143	16452	17595

Distribuições Bivariadas - Exemplo II

Total

	DRE+	DRE-	
PSA+	330	1047	1377
PSA-	1147	16952	18099
	1477	17999	19476

Distribuições Bivariadas - Exemplo II

Definimos eventos associados ao experimento:

A = DRE+	B = PSA+	C = Paciente tem câncer
A^c = DRE-	B^c = PSA-	C^c = Paciente não tem câncer

Na literatura, $P(A|C)$ é chamado de *sensibilidade* do DRE, e $P(A^c|C^c)$ é chadado de *especificidade* do DRE. Análogamente $P(B|C)$ e $P(B^c|C^c)$ são a sensibilidade e especificidade do PSA.

$P(C) = 1881/19476 = 0{,}0966$ é a *prevalência* do câncer de próstata em homens com 50 anos ou mais de idade.

Distribuições Bivariadas - Exercício

Se sortearmos um indivíduo da amostra, calcule as probabilidades dos seguintes eventos:

$$
\begin{array}{lll}
P(A|B) & P(A \cap B|C) & P(A \cup C) \\
P(A \cap B) & P(A \cup B) & P(B \cup C) \\
P(A \cap B \cap C) & P(A \cup B \cup C) & P(A \cup C|B)
\end{array}
$$