

Instituto de Matemática e Estatística
Universidade de São Paulo



Lista 6

MAE0330 - Análise Multivariada de Dados

Prof^a Lucia Pereira Barroso

Bruno Groper Morbin - n^oUSP 11809875

Luigi Pavarini de Lima - n^oUSP 11844642

São Paulo
05 novembro, 2022

1 Dados

Tabela 1.1: Conjunto de Dados - ‘mochila.xls’

	TM	CM	EM		TM	CM	EM		TM	CM	EM		TM	CM	EM		TM	CM	EM
1	C	4	D	47	E	1	D	93	L	3	D	139	E	1	E	185	C	5	E
2	C	4	E	48	E	1	D	94	L	3	D	140	E	6	D	186	C	4	D
3	C	4	D	49	E	2	E	95	E	1	D	141	E	1	E	187	E	1	E
4	C	5	E	50	C	4	E	96	E	1	D	142	E	1	E	188	E	1	E
5	C	4	D	51	E	1	E	97	E	1	D	143	L	3	D	189	E	1	E
6	C	4	E	52	E	1	E	98	E	1	D	144	C	4	D	190	E	1	D
7	C	4	D	53	C	4	E	99	C	5	E	145	E	1	E	191	C	4	D
8	E	1	D	54	E	1	E	100	C	4	D	146	L	3	D	192	E	1	D
9	C	4	D	55	C	4	E	101	L	3	D	147	E	1	E	193	C	4	D
10	L	3	E	56	C	4	D	102	E	1	E	148	E	1	E	194	E	1	E
11	C	4	D	57	E	1	E	103	C	5	E	149	E	1	D	195	E	1	E
12	C	4	D	58	E	1	E	104	E	1	E	150	E	1	D	196	E	1	E
13	C	4	D	59	C	5	E	105	C	4	A	151	C	5	E	197	E	1	E
14	C	4	D	60	L	3	A	106	E	1	D	152	C	4	D	198	E	1	D
15	C	4	E	61	C	5	A	107	E	1	E	153	C	4	E	199	E	1	D
16	C	4	E	62	E	1	D	108	E	1	A	154	C	4	E	200	E	1	D
17	C	4	D	63	E	1	E	109	E	1	E	155	E	1	E	201	E	1	A
18	E	1	D	64	E	1	D	110	L	3	A	156	E	2	E	202	C	5	E
19	E	6	D	65	E	1	D	111	C	5	E	157	C	4	E	203	C	4	D
20	L	3	D	66	C	5	D	112	E	1	A	158	E	1	E	204	E	1	E
21	E	1	A	67	C	4	E	113	E	1	E	159	E	1	E	205	C	4	E
22	E	1	A	68	E	1	E	114	E	1	D	160	E	1	D	206	E	1	E
23	C	4	D	69	C	4	D	115	C	4	D	161	C	4	E	207	E	1	E
24	L	3	E	70	C	4	D	116	E	1	E	162	E	1	E	208	E	1	E
25	E	1	A	71	C	4	D	117	E	1	E	163	C	4	E	209	E	1	E
26	L	3	E	72	L	3	D	118	L	3	E	164	E	1	E	210	E	1	E
27	E	1	D	73	E	1	E	119	E	1	E	165	C	4	D	211	C	4	D
28	C	4	A	74	E	2	D	120	E	1	E	166	E	1	A	212	E	1	E
29	C	4	D	75	C	4	D	121	E	1	A	167	C	5	E	213	E	1	E
30	L	3	D	76	C	5	D	122	L	3	E	168	E	3	E	214	E	1	E
31	L	3	D	77	E	1	E	123	L	6	E	169	L	2	E	215	L	3	E
32	C	4	E	78	C	4	A	124	E	1	E	170	E	6	E	216	E	2	E
33	C	5	E	79	E	1	D	125	E	1	E	171	E	1	E	217	C	4	E
34	C	4	D	80	C	4	A	126	E	1	A	172	C	4	A	218	E	1	E
35	C	4	D	81	E	1	D	127	E	1	E	173	E	1	E	219	E	1	E
36	E	1	E	82	C	4	E	128	L	3	E	174	E	1	E	220	E	1	E
37	C	4	E	83	C	4	E	129	E	4	A	175	E	1	E	221	E	1	E
38	C	4	E	84	E	1	D	130	E	1	A	176	C	4	D	222	E	1	D
39	E	1	D	85	E	1	D	131	C	4	A	177	E	1	D	223	L	3	E
40	E	1	A	86	E	1	A	132	E	1	E	178	C	4	D	224	L	3	E
41	E	1	E	87	E	1	A	133	C	5	D	179	E	1	E	225	E	1	E
42	C	5	E	88	C	4	E	134	E	1	E	180	C	4	D	226	E	1	E
43	E	1	E	89	E	6	A	135	E	1	A	181	E	1	D				
44	C	4	D	90	C	4	E	136	L	3	D	182	E	1	E				
45	E	1	E	91	C	4	A	137	C	4	E	183	C	4	D				
46	C	4	E	92	E	1	E	138	E	1	D	184	E	2	D				

TM: Tipo de mochila utilizada (E = escapular, L = lateral, C = carrinho)
CM: Modo de carregar a mochila (1 = nos dois ombros, 2 = em um ombro, 3 = no tronco, 4 = na mão direita, 5 = na mão esquerda, 6 = Outro)
EM: Lado da escoliose (D = direito, E = esquerdo, A = ausente)

2 Associações

2.1 EM & TM

Avalia-se a seguir a associação entre o lado da **escoliose** (EM) e o **tipo** de mochila (TM). De acordo com a Tabela 1.1, tem-se a seguinte tabela de contigência para EM e TM:

Tabela 2.1: Tabela de contigência EM & TM.

EM \ TM	C	E	L	Total
A	8	16	2	26
D	37	35	10	82
E	35	72	11	118
Total	80	123	23	226

2.2 EM & CM

Avalia-se a seguir a associação entre o lado da **escoliose** (EM) e o modo de **carregar** a mochila (CM).

Tabela 2.2: Tabela de contigência EM & CM.

CM \ EM	1	2	3	4	5	6	Total
A	14	0	2	8	1	1	26
D	31	2	10	34	3	2	82
E	67	4	10	24	11	2	118
Total	112	6	22	66	15	5	226

2.3 EM & TM & CM

Avalia-se a seguir a associação entre o lado da **escoliose** (EM), o **tipo** de mochila (TM) e o modo de **carregar** a mochila (CM).

```
> ## ANALISE DE CORRESPONDENCIA
>
>
>
> # tabela3 do texto
>
> tabela3 <- matrix(c(110, 102, 98, 61, 31, 122, 110, 98, 82, 42, 68, 66, 60, 82, 82, 24, 35, 48, 87, 92),
+   nrow = 5)
> tabela.SVD <- svd(tabela3)
> tabela.SVD
> tabela.SVD$d
>
>
> rownames(tabela3) <- c("A", "B", "C", "D", "E")
>
> colnames(tabela3) <- c("18-25", "26-40", "41-60", "+60")
>
> tabela3
>
> # tabela lida do Excel abrir planilha do Excel, iluminar e 'copiar'
> table <- read.table("clipboard")
> table
> rownames(table)
> colnames(table)
>
>
>
> # tabela lida de arquivo txt tabela <- read.table('arquivo.txt')
>
>
>
> ### matrizes de propor<U+FFFD><U+FFFD>o
>
>
>
> # porcentagem nas linhas
> tabelal <- prop.table(tabela3, 1)
> print(tabelal, digits = 3)
>
>
>
> # porcentagem nas colunas
> tabelac <- prop.table(tabela3, 2)
> print(tabelac, digits = 3)
>
>
>
> # porcentagem total
> tabelap <- prop.table(tabela3)
> tabelap
> print(tabelap, digits = 3)
```

```

>
>
> ## qui-quadrado e inercia total
>
> tabela.rowsum <- apply(tabela3, 1, sum)
> tabela.rowsum
> tabela.colsum <- apply(tabela3, 2, sum)
> tabela.colsum
> tabela.sum <- sum(tabela3)
> tabela.sum
> tabela.esp <- tabela.rowsum %/% tabela.colsum/tabela.sum
> tabela.esp
> chi <- (tabela3 - tabela.esp)^2/tabela.esp
> chi
>
> chi2 <- sum(chi)
> chi2
>
> inercia <- chi2/tabela.sum
> inercia
>
>
>
> ## Análise de correspondência simples
>
> library(ca)
>
> help(ca)
>
> ac <- ca(tabela3, nd = 3)
> ac
>
> ac <- ca(tabela3)
>
> ac
> # mostra inercia principal e suas porcentagens massa distância qui-quadrado do centroide
+ # inercia coordenadas padrão (vértices)
>
> summary(ac)
>
> # autovalores = inércia principal nas tabelas de linhas e colunas, as quantidades são
+ # multiplicadas por 1000 massa, qualidade da representação e inércia de cada
+ # ponto k=1 e k=2 são as coordenadas principais (para mapa simétrico) cor
+ # contribui<U+FFFD><U+FFFD>o do eixo para a inércia do ponto ctr <U+FFFD>
+ # contribui<U+FFFD><U+FFFD>o do ponto para inércia do eixo
>
> # mapa simétrico
> plot(ac)
>
> # mapas assimétricos
> plot(ac, map = "rowprincipal")
> plot(ac, map = "colprincipal")
>
>
> # para incluir pontos suplementares, indicar os números das linhas ou colunas em que esses
+ # pontos estão na tabela, por exemplo se forem a linha 4 e a coluna 3
>
> ac <- ca(tabela3, suprow = 4, supcol = 3)
>
>
>

```

```
>
> # An<U+FFFD>lise de correspond<U+FFFD>ncia m<U+FFFD>ltipla
>
> # sa<U+FFFD>de versus idade
> idade <- read.table("clipboard")
> idade
> rownames(idade)
> colnames(idade)
>
>
>
> acI <- ca(idade)
> acI
> summary(acI)
> plot(acI)
>
>
>
> # sa<U+FFFD>de versus sexo
> sexo <- read.table("clipboard")
> sexo
> rownames(sexo)
> colnames(sexo)
>
>
>
> acS <- ca(sexo, nd = 1)
> acS
> summary(acS)
> plot(acS)
>
>
>
> # sa<U+FFFD>de versus idade e sexo
>
> # tabelas combinadas
>
> idasec <- read.table("clipboard")
> idasec
> rownames(idasec)
> colnames(idasec)
>
>
>
> acISC <- ca(idasec)
> acISC
> summary(acISC)
> plot(acISC)
>
>
>
> # tabelas empilhadas
>
> idasee <- read.table("clipboard")
> idasee
> rownames(idasee)
> colnames(idasee)
>
>
>
> acISE <- ca(idasee)
> acISE
> summary(acISE)
> plot(acISE)
>
>
>
> # MCA - usando tabela indicadora dados casamento: copiar somente colunas B a E da planilha
+ # indicador
```

```
>
> casamento <- read.table("clipboard")
> casamento
>
>
>
> acmI <- mjca(casamento, lambda = "indicator")
> acmI
> summary(acmI)
> plot(acmI)
> plot(acmI, map = "rowprincipal")
> plot(acmI, map = "colprincipal")
>
> # se a leitura do banco de dados for da coluna A a E
> acmB <- mjca(casamento[, 2:5], lambda = "Burt")
> acmB
> summary(acmB)
> plot(acmB)
>
>
> acmJ <- mjca(casamento[, 2:5], lambda = "JCA")
> acmJ
> summary(acmJ)
> plot(acmJ)
```