



Universidade Estadual de Santa Cruz
Departamento de Ciências Exatas e Tecnológicas - DCET
Curso: Engenharia Elétrica
Discente: José Cláudio Faria
Docente: Sarah Modesto Sanches(202012319) e Kevin Moura(201811019).

Análise Exploratória de dados: CET173 - Probabilidade e Estatística

Matriculas: 202012319 201811019 201811019
Os dados foram simulados computacionalmente e tem finalidades apenas acadêmicas e didáticas para fins de avaliação

11 de novembro de 2021 ILHÉUS-BA

1 AED: Apresentações tabulares e gráficas (2.0)

1.1 (1.0) Diagrama de caixa (boxplot) para Y1 e Y2:

1.1.1 (0.5) Antes e após a eliminação de possíveis outliers - sem distinção de sexo:

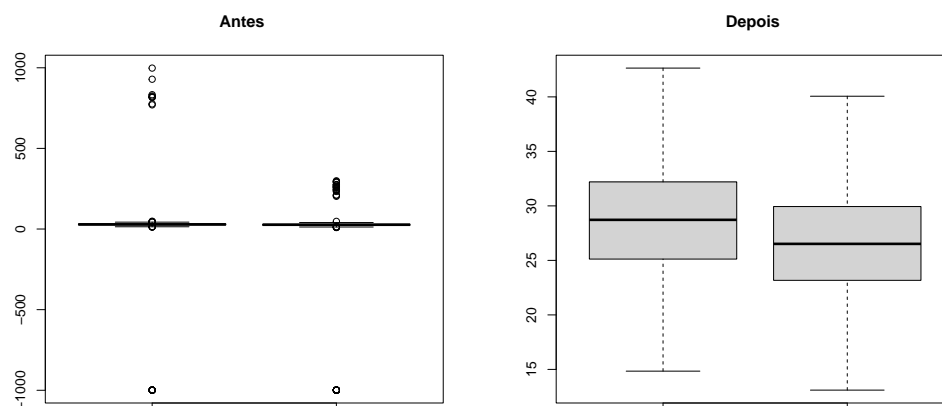


Figura 1: Diagrama de caixa de Y1 (un) e Y2 (un) antes e após a eliminação de outliers, UESC/BA - 2021.

1.1.2. (0.5) Após a eliminação de possíveis outliers para cada sexo:

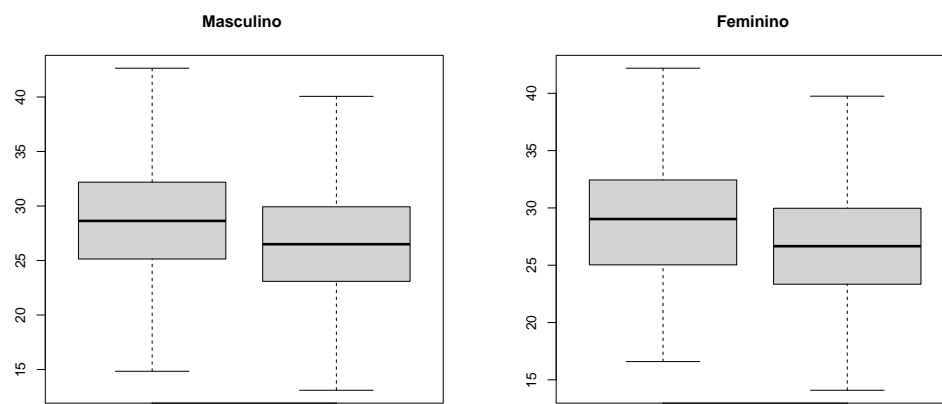


Figura 2: Diagrama de caixa de Y1 (un) e Y2 (un) (sexo masculino e feminino, respectivamente, UESC/BA - 2021.

1.2 Para Y1

1.2.1 (0.5)Apresentações tabulares

Tabela 1: Tabela de distribuição de frequência de Y1 (sexo masculino), UESC/BA - 2021

	Class limits	f	rf	rf(%)	cf	cf(%)
1	[14.692,17.056)	14.00	0.01	0.94	14.00	0.94
2	[17.056,19.421)	32.00	0.02	2.14	46.00	3.08
3	[19.421,21.785)	95.00	0.06	6.36	141.00	9.44
4	[21.785,24.15)	163.00	0.11	10.91	304.00	20.35
5	[24.15,26.514)	211.00	0.14	14.12	515.00	34.47
6	[26.514,28.879)	263.00	0.18	17.60	778.00	52.07
7	[28.879,31.244)	247.00	0.17	16.53	1025.00	68.61
8	[31.244,33.608)	206.00	0.14	13.79	1231.00	82.40
9	[33.608,35.973)	146.00	0.10	9.77	1377.00	92.17
10	[35.973,38.337)	85.00	0.06	5.69	1462.00	97.86
11	[38.337,40.702)	22.00	0.01	1.47	1484.00	99.33
12	[40.702,43.066)	10.00	0.01	0.67	1494.00	100.00

Tabela 2: Tabela de distribuição de frequência de Y1 (sexo feminino), UESC/BA - 2021

	Class limits	f	rf	rf(%)	cf	cf(%)
1	[16.424,19.043)	12.00	0.03	3.17	12.00	3.17
2	[19.043,21.662)	25.00	0.07	6.61	37.00	9.79
3	[21.662,24.28)	41.00	0.11	10.85	78.00	20.63
4	[24.28,26.899)	62.00	0.16	16.40	140.00	37.04
5	[26.899,29.518)	62.00	0.16	16.40	202.00	53.44
6	[29.518,32.137)	76.00	0.20	20.11	278.00	73.54
7	[32.137,34.756)	46.00	0.12	12.17	324.00	85.71
8	[34.756,37.374)	35.00	0.09	9.26	359.00	94.97
9	[37.374,39.993)	16.00	0.04	4.23	375.00	99.21
10	[39.993,42.612)	3.00	0.01	0.79	378.00	100.00

1.2.2 (0.5) Histograma e o polígono de frequência acumulada

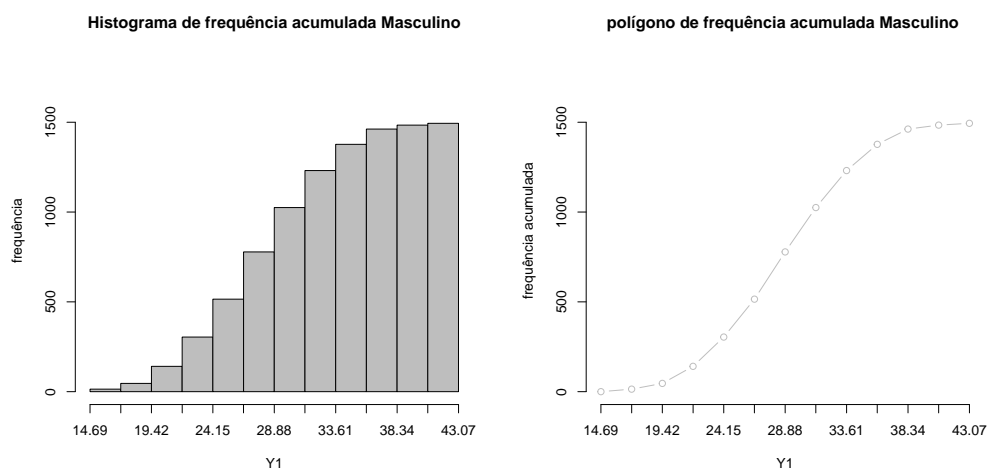


Figura 3: Histograma e polígono de frequência acumulada de Y1 (un) (sexo masculino), UESC/BA - 2021.

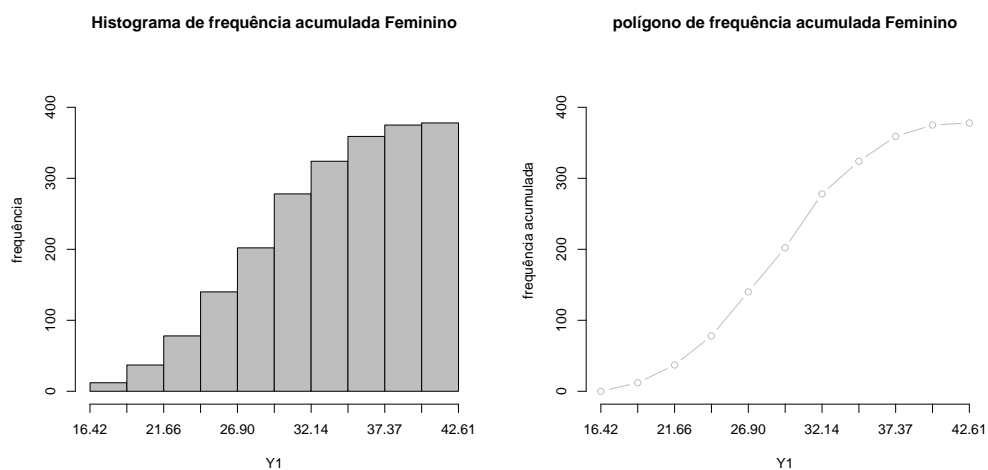


Figura 4: Histograma e polígono de frequência acumulada de Y1 (un) (sexo feminino), UESC/BA - 2021.

2 AED: medidas estatísticas básicas(3.0).

2.1 (1.5) AED: Medidas determinadas a partir dos vetores

2.1.1 (0.5) Tendência central

Tabela 3: Medidas de tendência central (sexo masculino), UESC/BA - 2021

var	n	m	md
Y1	28.66	27.81	28.63
Y2	26.43	31.01	26.49

Tabela 4: Medidas de tendência central (sexo feminino), UESC/BA - 2021

var	n	m	md
Y1	28.86	33.15	29.03
Y2	26.52	27.14	26.66

2.1.2 (0.5) Posição

Tabela 5: Quartis dos usuários (sexo masculino), UESC/BA - 2021

	Y1	Y2
25%	25.14	23.09
50%	28.63	26.49
75%	32.19	29.93

Tabela 6: Quartis dos usuários (sexo feminino), UESC/BA - 2021

	Y1	Y2
25%	25.04	23.36
50%	29.03	26.66
75%	32.43	29.96

Tabela 7: Decis dos usuários (sexo Masculino), UESC/BA - 2021

	Y1	Y2
10%	21.97	20.00
20%	24.09	22.10
30%	25.91	23.77
40%	27.37	25.26
50%	28.63	26.49
60%	29.96	27.71
70%	31.41	29.12
80%	33.10	30.79
90%	35.46	32.64

Tabela 8: Decis dos usuários (sexo Feminino), UESC/BA - 2021

	Y1	Y2
10%	21.97	20.00
20%	24.09	22.10
30%	25.91	23.77
40%	27.37	25.26
50%	28.63	26.49
60%	29.96	27.71
70%	31.41	29.12
80%	33.10	30.79
90%	35.46	32.64

2.1.3 (0.5) Dispersão

Tabela 9: Dispersão dos usuários (sexo masculino), UESC/BA - 2020

	a.t	variância	d.padrão	c.v
1	27.80	25.56	5.06	17.64
2	26.96	23.34	4.83	18.27

Tabela 10: Dispersão dos usuários (sexo feminino), UESC/BA - 2020

	a.t	variância	d.padrão	c.v
1	25.60	26.79	5.18	17.94
2	25.66	25.53	5.05	19.05

2.2 AED: Medidas determinadas a partir de apresentações tabulares (1.5)

Tabela 11: Tabela de distribuição de frequência reconstruída de publicação, UESC/BA - 2020

	Classes	f	rf(%)	cf(%)
1	[10, 020)	9.00	3.90	3.90
2	[20, 030)	18.00	7.79	11.69
3	[30, 040)	28.00	12.12	23.81
4	[40, 050)	39.00	16.88	40.69
5	[50, 060)	47.00	20.35	61.04
6	[60, 070)	39.00	16.88	77.92
7	[70, 080)	27.00	11.69	89.61
8	[80, 090)	17.00	7.36	96.97
9	[90, 100)	7.00	3.03	100.00

2.3 (0.5) Tendência central

Tabela 12: Medidas de tendência central, UESC/BA - 2021

	var	n	m	md
medidas	54.44	55.00	54.57	

2.1.2 (0.5) Posição

Tabela 13: Quartis dos usuários, UESC/BA - 2021

	10%	20%	30%
medidas	40.71	54.57	68.27

Tabela 14: Decis dos usuários, UESC/BA - 2021

	10%	20%	30%	40%	50%	60%	70%	80%	90%
medidas	27.83	36.86	43.67	49.59	54.57	59.49	65.31	71.78	80.53

2.4 (0.5) Dispersão

Tabela 15: Dispersão dos usuários, UESC/BA - 2020

	variância	d.padrão	c.v
1	377.51	19.43	2.80

3 AED: Medidas estatísticas de associação e regressão linear (4.0)

3.1 (1.5) Associação

3.1.1 (0.5) Estimativas: covariância e correlação linear simples

Tabela 16: Matriz de variâncias e covariâncias (sexo masculino), UESC/BA - 2021

	Y1	Y2
Y1	25.56	23.00
Y2	23.00	23.34

Tabela 17: Matriz de variâncias e covariâncias (sexo feminino), UESC/BA - 2021

	Y1	Y2
Y1	26.79	-20.63
Y2	-20.63	25.53

Tabela 18: Matriz de correlações lineares simples (sexo masculino), UESC/BA - 2021

	Y1	Y2
Y1	1.00	-0.79
Y2	-0.79	1.00

Tabela 19: Matriz de correlações lineares simples (sexo feminino), UESC/BA - 2021

	Y1	Y2
Y1	1.00	-0.79
Y2	-0.79	1.00

3.1.2 (0.5) Diagrama de dispersão dos dados

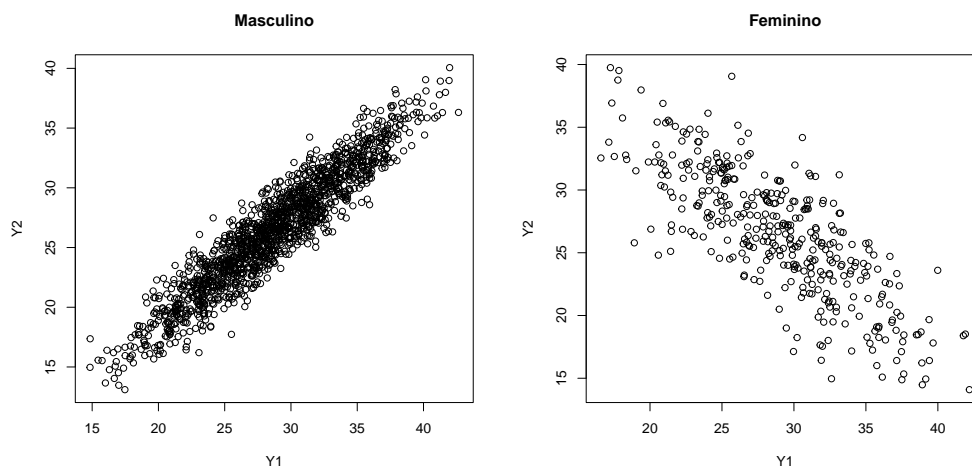


Figura 5: Diagrama de dispersão de Y1 (un) e Y2 (un) (sexo masculino e feminino, respectivamente), UESC/BA - 2021.

3.1.3 (0.5) Comparação de estudos semelhantes

3.2 3.2 Regressão linear (2.5)

3.2.1 (1.0) Ajustamento

Tabela 20: Polinômio grau I, UESC/BA - 2021

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.7291	0.6237	2.77	0.0242
x	1.3072	0.1052	12.43	0.0000

Tabela 21: Polinômio grau II, UESC/BA - 2021

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.6429	0.6431	1.00	0.3508
x	2.0405	0.2995	6.81	0.0003
I(x ²)	-0.0733	0.0288	-2.54	0.0385

3.2.2 (0.5) Diagrama de dispersão com modelos ajustados

```
## Error in eval(predvars, data, env):  objeto 'y' não encontrado
## Error in eval(predvars, data, env):  objeto 'y' não encontrado
```

Figura 6: Diagrama de dispersão dos dados, modelos ajustados e respectivos r^2 , UESC/BA - 2021.

3.2.3 (0.5) Qual modelo melhor explica o fenômeno em estudo?

3.2.4 (0.5) Critérios de ajustamento e escolha de modelos

3.3 Contextualização (1.0))