

Universidade Estadual de Santa Cruz - UESC

Departamento de Ciências Exatas e Tecnológicas - DCET

CET083 - Probabilidade e Estatística

Curso de Ciência da Computação

Prof. José Cláudio Faria

Correção computacional usando R (<http://www.r-project.org/>)

`Matriculas = 202121141 - 202121148`

Os dados foram simulados computacionalmente e tem finalidades apenas acadêmicas e didáticas para fins de avaliação.

AED: Apresentações tabulares e gráficas

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

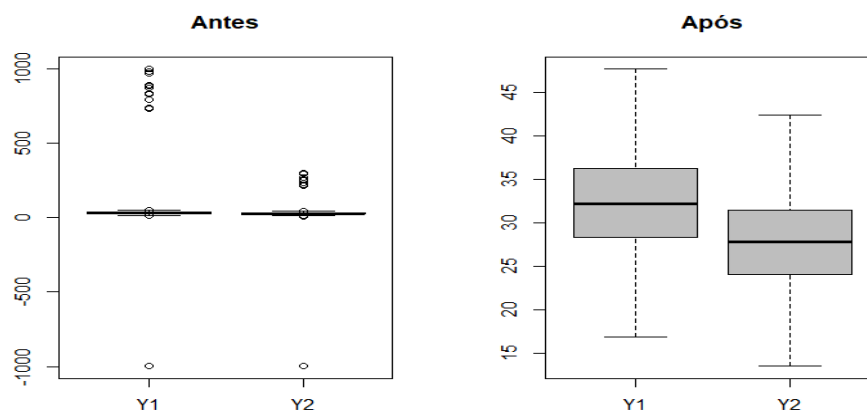


Figura 1 – Diagrama de caixa de Y1 (*un*) e Y2 (*un*) antes e após a eliminação de outliers, UESC/BA - 2023.

1.1.2 (0.5) Após a eliminação de possíveis outliers - com distinção de sexo

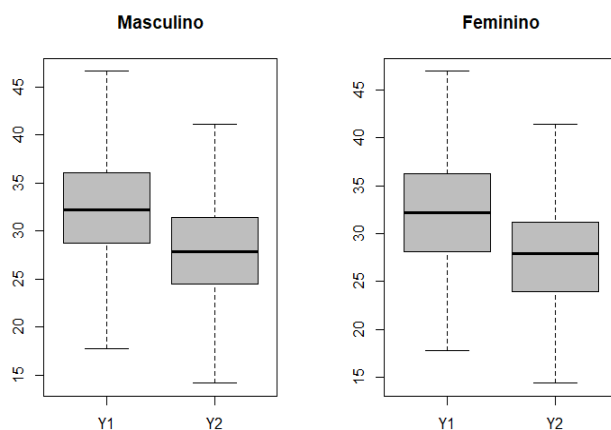


Figura 2 – Diagrama de caixa de Y1 (*un*) e Y2 (*un*) (sexo masculino e feminino, respectivamente), UESC/BA - 2023.

I.2 (1.0) Para Y1

I.2.1 (0.5) Apresentações tabulares

Tabela 1 – Tabela de distribuição de frequência de Y1 (*un*) (sexo masculino), UESC/BA – 2023

class limits	f	rf	rf(%)	cf	cf(%)
[16.75,19.37)	18	0.01	1.28	18	1.28
[19.37,21.99)	39	0.03	2.77	57	4.05
[21.99,24.61)	61	0.04	4.34	118	8.39
[24.61,27.23)	129	0.09	9.17	247	17.57
[27.23,29.85)	225	0.16	16.00	472	33.57
[29.85,32.47)	283	0.20	20.13	755	53.70
[32.47,35.09)	225	0.16	16.00	980	69.70
[35.09,37.72)	215	0.15	15.29	1195	84.99
[37.72,40.34)	127	0.09	9.03	1322	94.03
[40.34,42.96)	53	0.04	3.77	1375	97.80
[42.96,45.58)	22	0.02	1.56	1397	99.36
[45.58,48.2)	9	0.01	0.64	1406	100.00

Tabela 2 – Tabela de distribuição de frequência de Y1 (*un*) (sexo feminino), UESC/BA – 2023

class limits	f	rf	rf(%)	cf	cf(%)
[17.563,20.276)	14	0.02	2.36	14	2.36
[20.276,22.989)	27	0.05	4.55	41	6.90
[22.989,25.703)	46	0.08	7.74	87	14.65
[25.703,28.416)	70	0.12	11.78	157	26.43
[28.416,31.129)	97	0.16	16.33	254	42.76
[31.129,33.843)	107	0.18	18.01	361	60.77
[33.843,36.556)	96	0.16	16.16	457	76.94
[36.556,39.269)	70	0.12	11.78	527	88.72
[39.269,41.983)	39	0.07	6.57	566	95.29
[41.983,44.696)	21	0.04	3.54	587	98.82
[44.696,47.409)	7	0.01	1.18	594	100.00

1.2.2 (0.5) Histograma e polígono de frequência acumulada

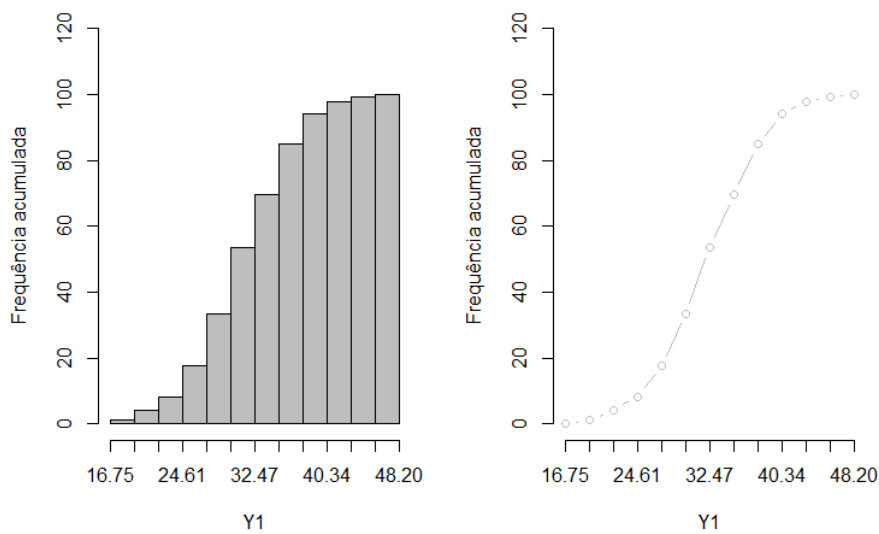
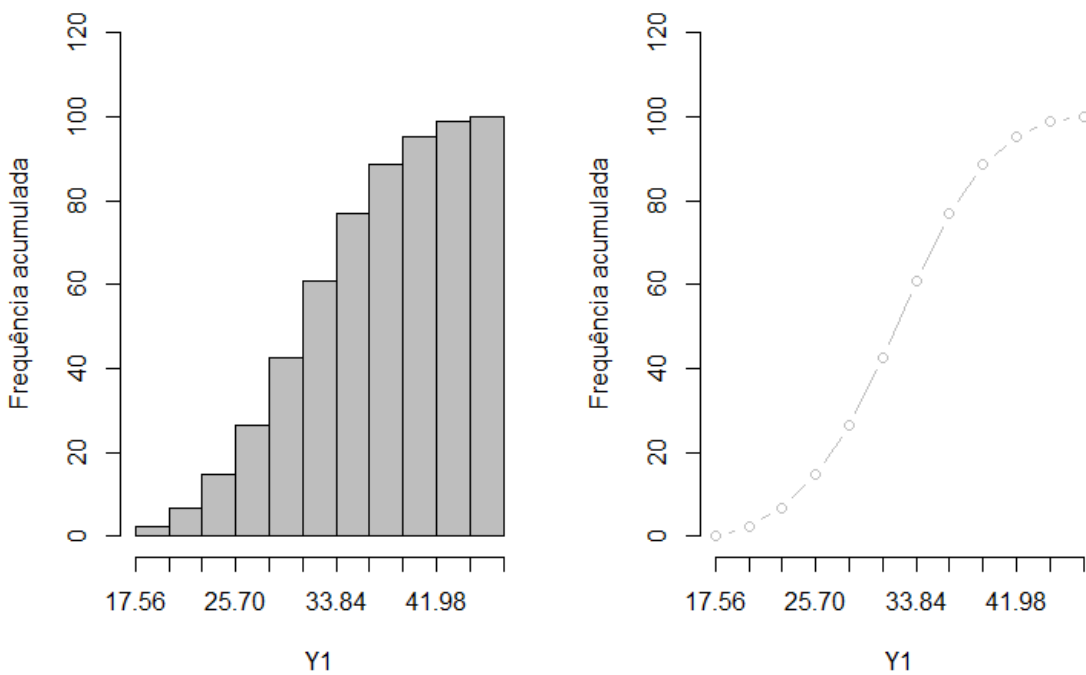


Figura 3 – Histograma e polígono de frequência acumulada de Y1 (un) (sexo masculino), UESC/BA - 2023.



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 – 2023

	n	m	md
Y1	32.14806	32.20	40
Y2	27.88495	27.88	33

Tabela 4 – Medidas de tendência central (sexo feminino), UESC/BA – 2023

	n	m	md
Y1	32.13113	32.20	11
Y2	27.73128	27.88	17

2.1.2 (0.5) Posição

Tabela 5 – Quartis dos usuários (sexo masculino), UESC/BA - 2023

	25%	50%	75%
Y1	28.6825	32.20	36.045
Y2	24.4850	27.88	31.415

Tabela 6 – Quartis dos usuários (sexo feminino), UESC/BA – 2023

	25%	50%	75%
Y1	28.09	32.20	36.2125
Y2	23.96	27.88	31.1975

Tabela 7 – Decis dos usuários (sexo masculino), UESC/BA – 2023

	10%	20%	30%	40%	50%	60%	70%	80%	90%
Y1	25.300	27.80	29.395	30.76	32.20	33.57	35.135	36.84	39.000
Y2	21.345	23.68	25.310	26.68	27.88	29.20	30.680	32.31	34.365

Tabela 8 – Decis dos usuários (sexo feminino), UESC/BA – 2023

	10%	20%	30%	40%	50%	60%	70%	80%	90%
Y1	24.291	26.936	29.078	30.690	32.20	33.558	35.212	37.178	39.525
Y2	20.603	23.142	24.708	26.316	27.88	28.838	30.394	32.156	34.918

2.1.3 (0.5) Dispersão

Tabela 9 – Dispersão dos usuários (sexo masculino), UESC/BA – 2023

	a.t	variância	d.padrão	c.v
Y1	30.80	29.36864	5.419284	16.85727
Y2	28.96	25.44752	5.044554	18.09060

Tabela 10 – Dispersão dos usuários (sexo feminino), UESC/BA – 2023

	a.t	variância	d.padrão	c.v
Y1	29.20	33.31210	5.771664	17.96284
Y2	28.03	28.80493	5.367023	19.35368

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

Tabela 11 – Tabela de distribuição de frequência reconstruída de publicação, UESC/BA – 2023

Classes	f	rf(%)	cf(%)
1 [10, 020)	9	3.85	3.85
2 [20, 030)	19	8.12	11.97
3 [30, 040)	27	11.54	23.50
4 [40, 050)	38	16.24	39.74
5 [50, 060)	49	20.94	60.68
6 [60, 070)	38	16.24	76.92
7 [70, 080)	27	11.54	88.46
8 [80, 090)	18	7.69	96.15
9 [90, 100)	9	3.85	100.00

2.2.1 (0.5) Tendência central

Tabela 12 – Medidas de tendência central

	m	md	mo
medida	54.87179	54.89796	55

2.2.2 (0.5) Posição

Tabela 13 – Medidas de posição: quartis

	25%	50%	75%
quartil	41.2963	54.89796	68.81579

Tabela 14 – Medidas de posição: decis

	10%	20%	30%	40%	50%	60%	70%	80%	90%
decil	27.57895	36.96296	44	50.15789	54.89796	59.67347	65.73684	71.89474	81.33333

2.2.3 (0.5) Dispersão

Tabela 15 – Medidas de dispersão

	amplitude	variância	d.padrão	c.v
medida	80	390.1545	19.75233	35.99724

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 – 2023

	Y1	Y2
Y1	29.36864	25.71211
Y2	25.71211	25.44752

Tabela 17 – Matriz de variâncias e covariâncias (sexo feminino), UESC/BA – 2023

	Y1	Y2
Y1	33.31210	-23.38178
Y2	-23.38178	28.80493

Tabela 18 – Matriz de correlações lineares simples (sexo masculino), UESC/BA – 2023

	Y1	Y2
Y1	1.0000000	0.9405307
Y2	0.9405307	1.0000000

Tabela 19 – Matriz de correlações lineares simples (sexo feminino), UESC/BA – 2023

	Y1	Y2
Y1	1.0000000	-0.7548196
Y2	-0.7548196	1.0000000

3.1.2 (0.5) Diagrama de dispersão dos dados

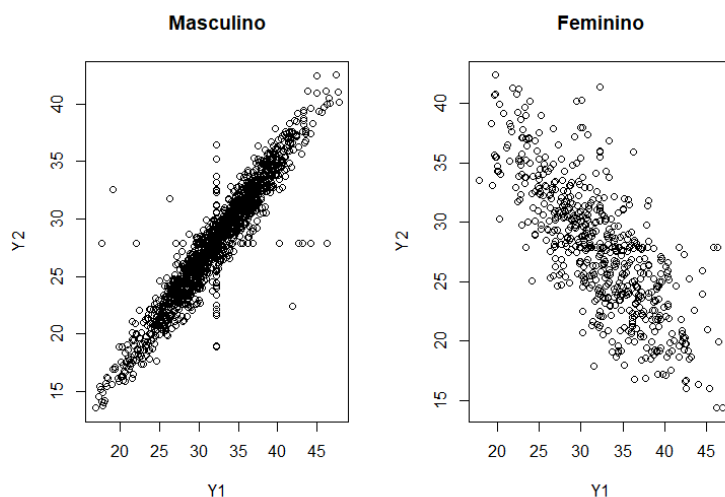


Figura 5 – Diagrama de dispersão de Y1 (un) e Y2 (un) (sexo masculino e feminino, respectivamente), UESC/BA – 2023

3.1.3 (0.5) Comparação de estudos semelhantes

Avaliar resposta e justificativa!

3.2 Regressão linear (2.5)

3.2.1 (1.0) Ajustamento

Tabela 20 – Polinômio grau I, UESC/BA – 2023

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.9643	0.7747	2.536	0.0349 *
x	1.2067	0.1306	9.240	1.53e-05 ***

Tabela 21 – Polinômio grau II, UESC/BA – 2023

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.23724	1.02771	1.204	0.26776
x	1.69750	0.47862	3.547	0.00939 **
I(x^2)	-0.04908	0.04608	-1.065	0.32217

3.2.2 (0.5) Diagrama de dispersão com modelos ajustados

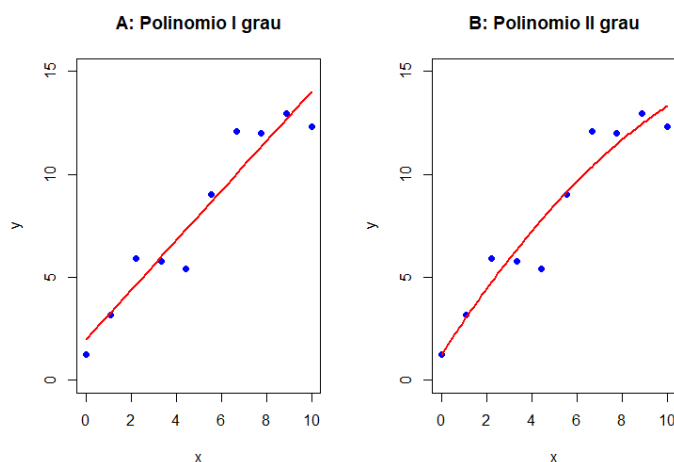


Figura 6 – Diagrama de dispersão dos dados, modelos ajustados e respectivos r^2 , UESC/BA - 2023.

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

O segundo modelo é o mais adequado para explicar os dados. No caso de um modelo quadrático, quanto maior for o valor de R^2 (coeficiente de determinação), mais explicativo é o modelo. Isso significa que ele fornece uma compreensão mais precisa da amostra, uma vez que o coeficiente de determinação é o principal critério utilizado para selecionar o melhor modelo. Além disso, o modelo quadrático também é preferível quando consideramos a orientação dos dados, pois ele mostra a tendência dos dados. Por exemplo, se a curva quadrática estiver voltada para baixo, indica que os dados estão diminuindo, enquanto se estiver voltada para cima, indica que os dados estão aumentando.

3.2.4 (0.5) Critérios de ajustamento e escolha de modelos. Avaliar resposta!

Não é apropriado fazer comparações diretas entre os modelos de ajuste estatístico devido à natureza controversa dessas questões. Por exemplo, ao ajustar uma regressão linear simples com intercepto zero a uma amostra de pares de dados usando R e Python, é possível encontrar diferenças que tornam os modelos controversos. Portanto, é importante considerar que existem variações e nuances nas abordagens estatísticas e de programação que podem levar a resultados divergentes, o que contribui para a falta de comparabilidade entre os modelos.

4 Contextualização (1.0) Avaliar resposta!

O artigo intitulado "ANÁLISE EXPLORATÓRIA DE DADOS DA CONDUTIVIDADE ELÉTRICA NAS ÁGUAS DO SISTEMA AQUÍFERO SERRA GERAL NO RIO GRANDE DO SUL E SANTA CATARINA" utiliza conceitos básicos da estatística descritiva e gráficos de distribuições amostrais para realizar o tratamento estatístico dos parâmetros. O método da Boxplot foi adotado para identificar valores anômalos, e o artigo destaca a vantagem desse método, que simplifica a classificação em poucas categorias com significados específicos.

A análise do artigo, considerando os temas abordados durante a disciplina, revela uma excelente utilização de representações tabulares. Além disso, o artigo faz bom uso de diagramas e de um modelo digital do terreno, o que contribui para uma compreensão mais aprofundada do conteúdo apresentado, utilizando uma linguagem técnica adequada.

Os autores do artigo são Marcos Alexandre de Freitas, Ari Roisenberg e José Leonardo Silva Andriotti, e foi publicado em 2016. O artigo pode ser acessado em:
<<https://rigeo.cprm.gov.br/xmlui/handle/doc/15584>>. Acesso em 03 de Dezembro de 2022.