

Nome: Deivid da Silva Galvão, Ra. 2408740

Para todas as questões abaixo, **interprete os resultados e apresente os códigos**. As bases de dados estão em anexo do google Classroom, já salvas em CSV, com o separador decimal em Inglês, ou seja, as decimais estão separadas por ponto.

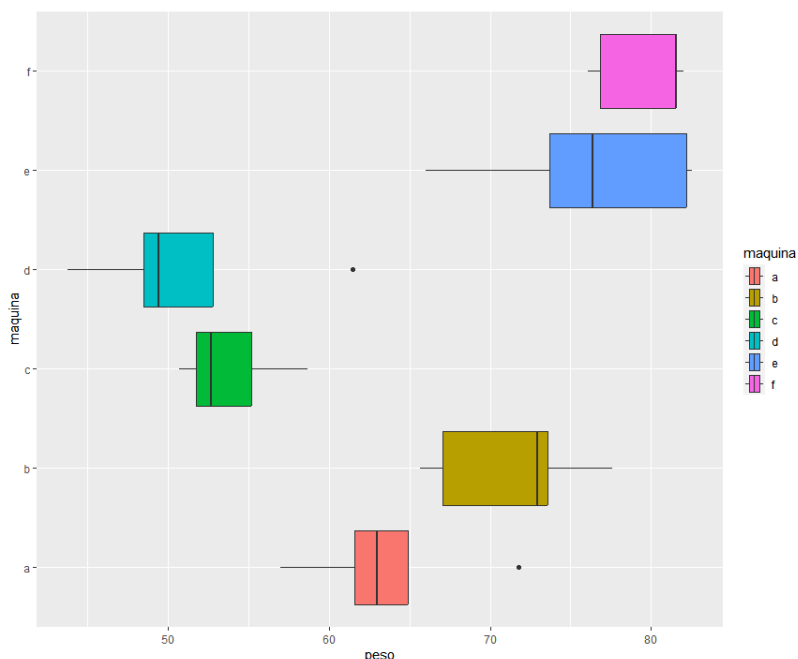
ex1) Qual o objetivo de uma análise de variância?

1 ponto

Resposta: O objetivo de uma análise de variância (ANOVA) é comparar 3 ou mais médias de grupos independentes verificando se há diferenças significativas entre elas, onde a partir dela pode-se concluir se há diferenças ou quais são melhores, ou piores para determinadas situações..

ex2) Um estudo foi elaborado para comparar o peso de vigas produzidas por 6 máquinas diferentes, sendo que quanto mais leve, melhores são as vigas. Para isso, 5 vigas de cada máquina foram testadas (dados1.csv). Faça uma análise de variância completa.
4 pontos

Resposta:



A Partir da análise do gráfico podemos ter uma ideia de a máquina D e a C produzem as vigas mais leves, porém nada podemos concluir efetivamente com essa análise e sem um teste mais completo da análise de variância.

Teste de Hipótese:

#H0: Todos são iguais

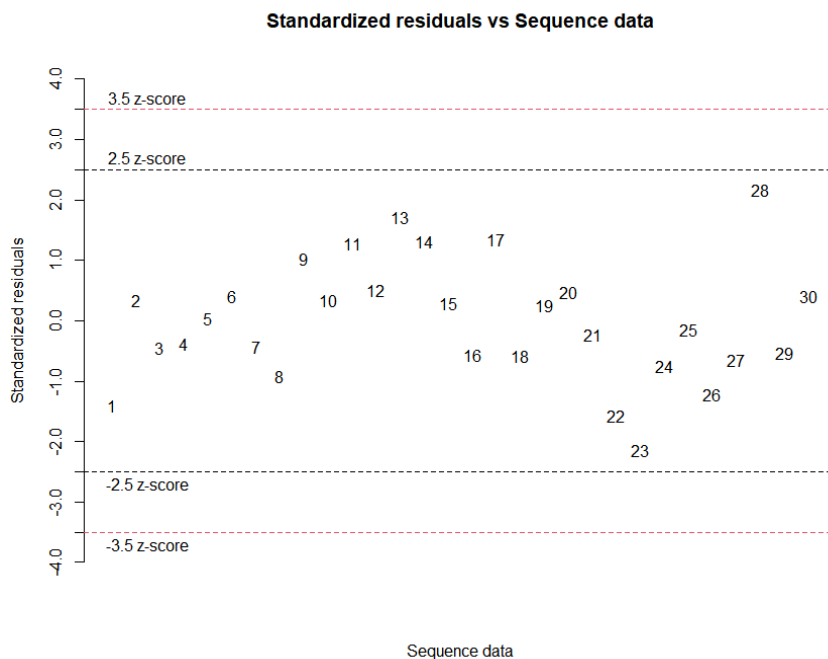
#H1: Ao menos um diferente

	df	type I SS	mean square	F value	p>F
treatments	5	3449.3293	689.8659	25.601	<0.001
Residuals	24	646.7232	26.9468	-	-

Podemos verificar que o p-valor está abaixo de 5% logo rejeita-se a hipótese nula e portanto ao menos 1 é diferente (H1), e portanto faz sentido verificar por meio do teste de Turkey (uma comparação par a par).

	treatment	mean	standard.error	tukey	snk	duncan	t	scott_knott
1	f	79.5973	2.3215	a	a	a	a	a
2	e	76.1673	2.3215	a	ab	ab	ab	a
3	b	71.3520	2.3215	ab	b	b	b	a
4	a	63.6562	2.3215	bc	c	c	c	b
5	c	53.8078	2.3215	cd	d	d	d	c
6	d	51.2072	2.3215	d	d	d	d	c

Podemos concluir que as máquinas “C” e “D” são iguais e são as que produzem as vigas mais leves.



vemos que a análise está correta, pois 95% dos pontos se encontram dentro do intervalo, logo o modelo está correto.

Códigos:

#ex2

install.packages('ggplot2')

library(ggplot2)

install.packages('easyanova')

library(easyanova)

dados1=read.csv('dados1.csv', sep = ',', dec='.', header = T)

```
ggplot(dados1,aes(x =peso, y=maquina, fill = maquina))+ geom_boxplot()
```

```
modelo=ea1(dados1, design = 1)
```

```
modelo$`Analysis of variance`
```

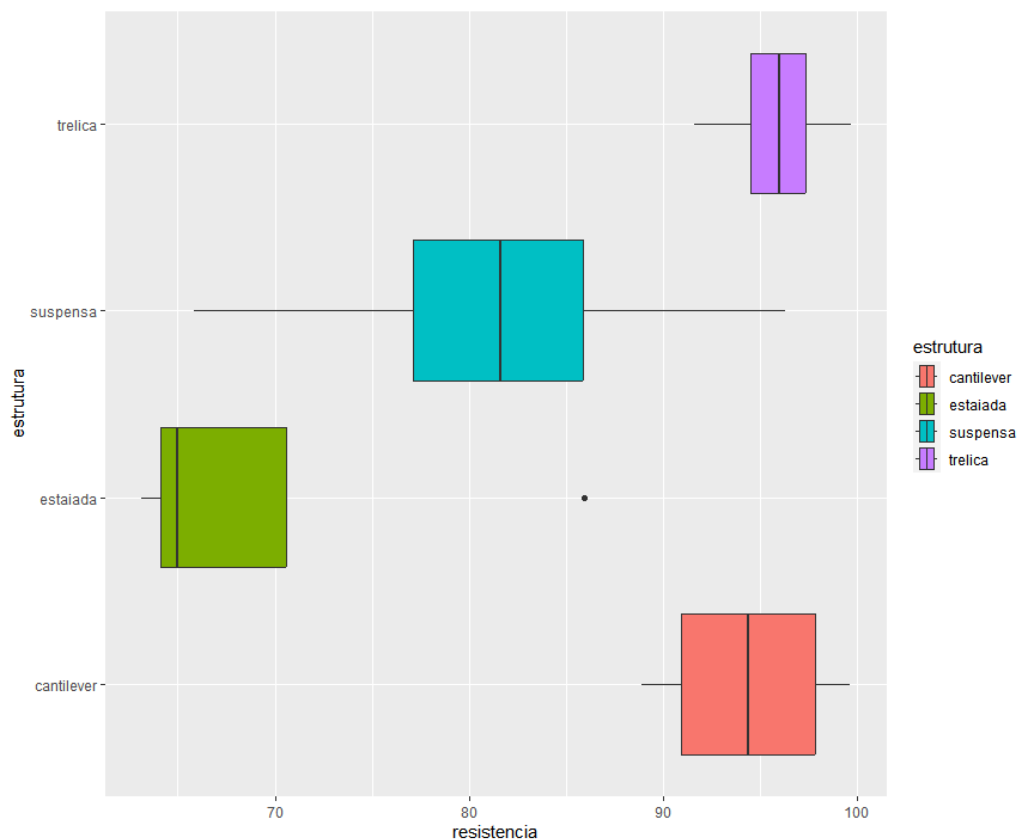
```
modelo$Means
```

ex3) Em um experimento, 4 tipos de estruturas de construção de pontes foram testadas, sendo elas suspensa, cantilever, estaiada e treliça. O objetivo do experimento foi verificar quais das estruturas possuem maior resistência, ou seja, quanto mais resistente, melhor. Como o espaço para a construção das mesmas não era homogêneo, os terrenos foram divididos em 4 tipos, sendo que em cada terreno as 4 estruturas foram testadas, totalizando 16 construções (dados2.csv). Faça uma análise de variância completa.

5 pontos

Resposta:

A Partir da análise do gráfico podemos ter uma ideia de que a estrutura treliça e a cantilever possuem as maiores resistências, porém nada podemos concluir efetivamente com essa análise e sem um teste mais completo da análise de variância.



Teste de Hipótese:

#H0: Todos são iguais

#H1: Ao menos um diferente

```

> anova(lm(y ~ treatment, data = dados))

```

	df	type III SS	mean square	F value	p>F
treatments	3	1797.1630	599.0543	6.0435	0.0154
blocks	3	31.9097	10.6366	0.1073	0.9537
residuals	9	892.1067	99.1230	-	-

Podemos verificar que o p-valor está abaixo de 5% logo rejeita-se a hipótese nula e portanto ao menos 1 é diferente (H1), e portanto faz sentido verificar por meio do teste de Turkey (uma comparação par a par).

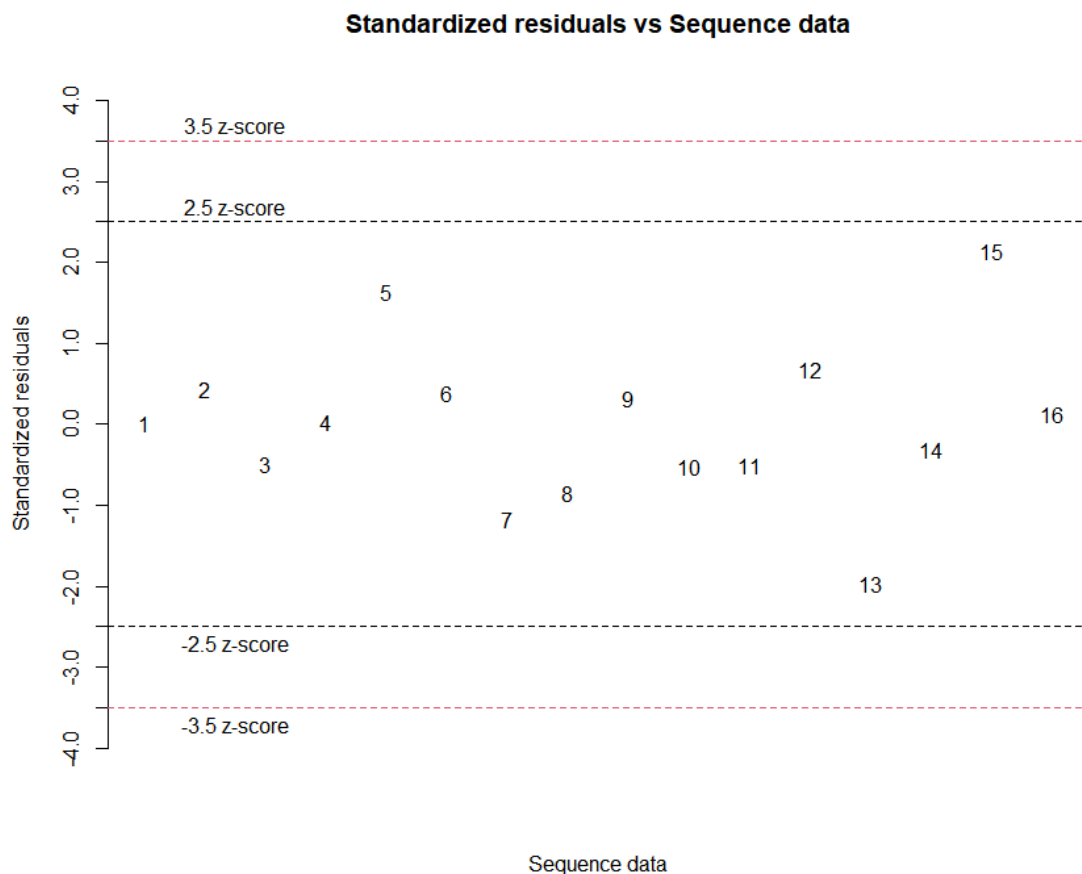
```

> TukeyHSD(lm(y ~ treatment, data = dados))

```

	treatment	adjusted.mean	standard.error	tukey	snk	duncan	t	scott_knott
1	trelica	95.8081	4.978	a	a	a	a	a
2	cantilever	94.3311	4.978	a	a	a	a	a
3	suspensa	81.3387	4.978	ab	ab	ab	ab	b
4	estaiada	69.7559	4.978	b	b	b	b	b

Podemos concluir que os tipos "trelica", "cantilever" e "suspensa" são iguais e são as que apresentam maiores resistências.



Por fim vemos que a análise está correta, pois 95% dos pontos se encontram dentro do intervalo, logo o modelo está correto.

Códigos: #ex3

`dados2=read.csv('dados2.csv', sep = ',', dec='.', header = T)`

`ggplot(dados2,aes(x = resistencia, y= estrutura, fill = estrutura))+ geom_boxplot()`

`modelo=eq1(dados2, design = 2)`

`modelo$`Analysis of variance``

`modelo$`Adjusted means``