

**UNIVERSIDADE DA REGIÃO DE JOINVILLE - UNIVILLE**

Bacharelado em Engenharia de Software (BES)

## **Estatística para computação**

**Professora Priscila Ferraz Franczak**

Engenheira Ambiental - UNIVILLE

Mestre em Ciência e Engenharia de Materiais - UDESC

Doutora em Ciência e Engenharia de Materiais - UDESC

[priscila.franczak@gmail.com](mailto:priscila.franczak@gmail.com)

# Plano de Aula

## Análise de variância - ANOVA (parte 2)

1. Teste de Tukey
2. Exercícios



- Os **testes de comparação de média** servem como um complemento para o estudo da análise de variância.
- Há vários testes de comparação de médias, entre os quais podemos citar: **teste de Tukey**, teste de Duncan, teste de Scheffé, teste de Dunnet e teste de Bonferroni.

# 1. Teste de Tukey

- É um dos testes de comparação de média mais utilizados, por ser bastante rigoroso e fácil aplicação;



- Também é conhecido como Teste de Tukey HSD (Teste de Tukey da Diferença Honestamente Significativa).
- O teste de Tukey foi desenvolvido por John Wilder Tukey e apresentado em 1949 no artigo titulado “*Comparing Individual Means in the Analysis of Variance*” (Biometrics. 5 (2): 99–114. [JSTOR 3001913](https://www.jstor.org/stable/3001913)).

- Quando os tamanhos amostrais dos grupos são iguais, o Teste de Tukey é um teste exato, ou seja, para o conjunto de todas as comparações par a par, a taxa de erro do conjunto dos testes é exatamente  $\alpha$  (nível de significância) e o intervalo de confiança é também exatamente  $1 - \alpha$ .



- É utilizado para testar toda e qualquer diferença entre duas médias de tratamento;
- É aplicado quando o teste “F” para tratamentos da ANOVA (análise de variância) for significativo (variâncias diferentes).

Para realizar o Teste de Tukey, deve ser levada em conta as seguintes suposições:

- As observações são independentes dentro e entre os grupos;
- Os grupos devem ser normalmente distribuídos;
- A variância dentro do grupo deve ser constante.



- A forma mais fácil de usar o teste de Tukey no R é empregando o comando `TukeyHSD( )`, do pacote `stats`, que já vem na instalação básica do R e é carregado sempre que o programa é iniciado.
- Para fazer o teste de Tukey usaremos os seguintes dados:

```
> porosidade<-c(25,31,22,33,  
+               26,25,26,29,  
+               20,28,28,31,  
+               23,27,25,34,  
+               21,24,29,28)  
> biocimento<-factor(rep(paste("bioc", 1:4, sep=""),5))  
> tabela<-data.frame(biocimento,porosidade)  
> tabela
```

	biocimento	porosidade
1	bioc1	25
2	bioc2	31
3	bioc3	22
4	bioc4	33
5	bioc1	26
6	bioc2	25
7	bioc3	26
8	bioc4	29
9	bioc1	20
10	bioc2	28
11	bioc3	28
12	bioc4	31
13	bioc1	23
14	bioc2	27
15	bioc3	25
16	bioc4	34
17	bioc1	21
18	bioc2	24
19	bioc3	29
20	bioc4	28



```
> ANOVA<-aov(porosidade~biocimento,tabela)
> resultado<-TukeyHSD(ANOVA,"biocimento")
> resultado
```

Tukey multiple comparisons of means  
95% family-wise confidence level

Fit: aov(formula = porosidade ~ biocimento, data = tabela)

```
$biocimento
```

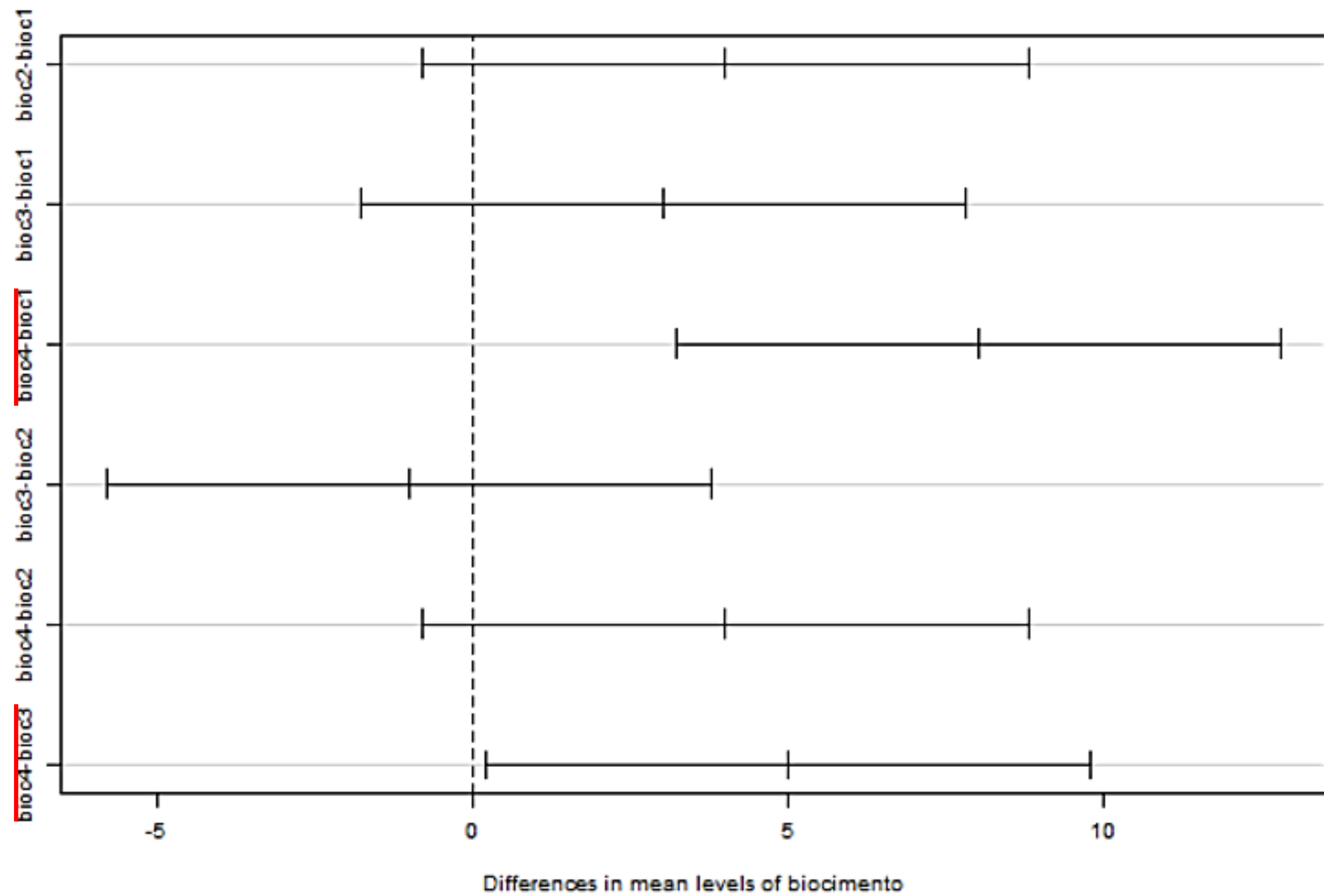
	diff	lwr	upr	p adj
bioc2-bioc1	4	-0.7874018	8.787402	0.1192178
bioc3-bioc1	3	-1.7874018	7.787402	0.3123298
bioc4-bioc1	8	3.2125982	12.787402	<u>0.0010547</u>
bioc3-bioc2	-1	-5.7874018	3.787402	0.9313122
bioc4-bioc2	4	-0.7874018	8.787402	0.1192178
bioc4-bioc3	5	0.2125982	9.787402	<u>0.0391175</u>

- Podemos visualizar os resultados graficamente:

```
plot(resultado, cex=0.6)
```

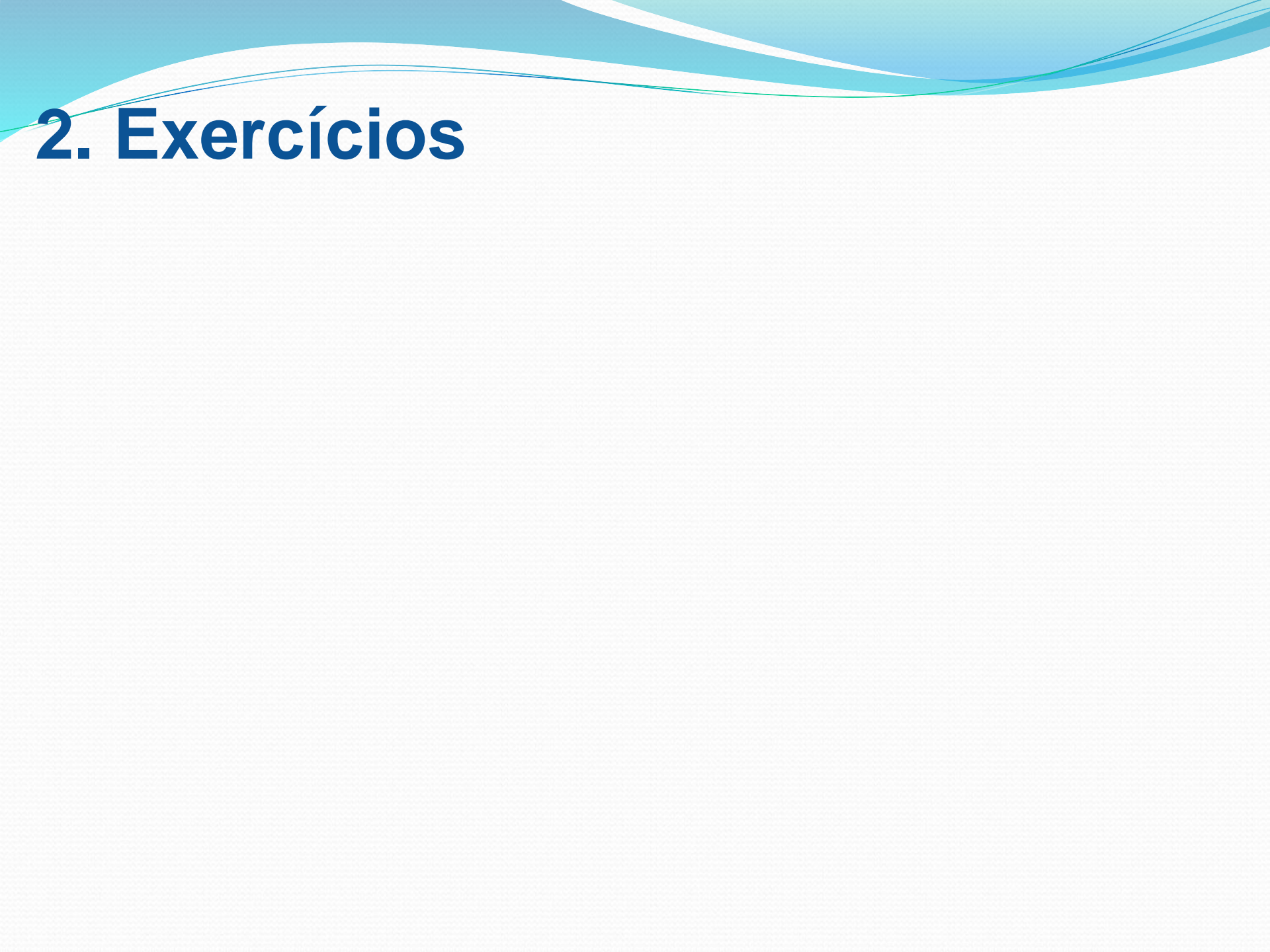


# 95% family-wise confidence level



- A interpretação dos resultados é a seguinte:
- Se o intervalo contiver o zero, a diferença entre as médias dos níveis do fator será não significativa para aquele par de médias em teste (médias consideradas iguais).
- Caso esse intervalo não contenha o zero, a diferença, portanto, será significativa ao nível de confiança em que o teste foi realizado.





# 2. Exercícios