



UNIVERSIDADE FEDERAL DE RONDÔNIA
CAMPUS DE JI-PARANÁ
DEPARTAMENTO DE ENGENHARIA AMBIENTAL



Estatística II


Prof.ª Renata Gonçalves Aguiar

1



Análise de Variância


2



ANOVA

A análise de variância (ou ANOVA, de *ANalysis Of VAriance*) é uma poderosa técnica estatística desenvolvida por R. A. Fisher e adaptada por Snedecor.

3




ANOVA

Extensão do teste *t*.

Compara qualquer número de médias.


4



ANOVA

Por que não fazer uma série de razões *t*?

5




ANOVA

Pearson (1942) mostrou que a probabilidade (P) de se cometer um erro do Tipo I aumenta com o número de médias que estão sendo comparadas ($\alpha = 0,05$).

$H_0: \mu_1 = \mu_2$	→	P = 0,05
$H_0: \mu_1 = \mu_2 = \mu_3$	→	P = 0,14
$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$	→	P = 0,26

6




ANOVA

Estatística II

A ANOVA mantém o erro do Tipo I constante

7

7



ANOVA

Estatística II

Experimentos inteiramente ao acaso, fator único.

Esse modelo que estudaremos é o mais simples de ANOVA.

8


8

Nota

No caso particular de um experimento com dois tratamentos, tanto se pode aplicar um teste t como a ANOVA.

9

9




Procedimentos para a ANOVA

Estatística II

A ANOVA só deve ser realizada se forem satisfeitas algumas pressuposições que serão discutidas posteriormente.

10

10




Procedimentos para a ANOVA

Estatística II

Antes de proceder às etapas da ANOVA, é importante analisar graficamente os dados de um experimento planejado.

11

11



Hipóteses


Estatística II

$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

H_1 : nem todas as médias das populações são iguais.

12

12




O experimento precisa ser bem delineado. No caso de experimentos inteiramente ao acaso, é essencial que as unidades experimentais utilizadas no experimento sejam, de início, similares e que a designação dos tratamentos às unidades tenha sido ao acaso.

13


13

Importantíssimo

Se isso não for feito não se deve concluir que os tratamentos são diferentes, mesmo diante de um teste F significante.




14



Planejamento


Os métodos estatísticos são essenciais para um bom experimento. Todos os experimentos são planejados. Infelizmente, alguns deles são pobremente planejados.



Fontes valiosas são usadas inefficientemente

15

15




Comparações Múltiplas

Estadística II

Teste de Tukey	→	Rigorous e fácil
Teste de Dunnett	→	Controle
Teste de Scheffé	→	Diferentes números de observações por tratamento

16

16




Situação-Problema 22

Estadística II

Utilize os dados médios de temperatura do ar da Reserva Biológica do Jaru nos meses de setembro, outubro e dezembro de 2013 e faça uma Anova (dados na página pessoal) em um programa estatístico. Qual conclusão chegaram?

17

17



Aplicação

Biogeosciences

Annual litterfall dynamics and nutrient deposition depending on elevation and land use at Mt. Kilimanjaro

J. Becker¹, H. Pabst¹, J. Mnyonga², and V. Kuznyakov^{1,3}

2015

18

18



UNIR

Estatística II

**Despertando o(a) o(a)
Discente Ativo(a)**

DEEA
Departamento de Engenharia Ambiental

19

19



UNIR

Estatística II

Discente Ativo

XIX SEMAT
PRATAS DA CASA:
DA GRADUAÇÃO À CARREIRA PROFISSIONAL
22 a 24 de Outubro de 2019

UNIR
FUNDAÇÃO UNIVERSITÁRIA
FÍSICA-SEMPERPARA

20

20



UNIR

Estatística II

Discente Ativo

Aproveitem as
oportunidades

Por R. G. Aguiar

21

21



UNIR

Estatística II

Discente Ativo

**XIII SEMANA
DA FÍSICA**

22

22



UNIR

Estatística II

Discente Ativo

Responda hoje o
questionário e
participe com sua
opinião

Foco na
melhoria

23

23



UNIR

Estatística II

Repor Aula

Dia 08.11.2019 às 8 h no LABGEO

24

24

Aula no Laboratório da Estatística

Dias 05, 08 e 12.11 – LABGEO

Trazer *notebook*

Ativar a Análise de Dados

Instalar o BioEstat 5.0
www.mamiraua.org.br

25

25

Aula no Laboratório da Estatística


Dias 05, 08 e 12.11 – LABGEO

Trazer os dados do trabalho na aula do dia 12.11.

Publicarei os arquivos das aulas práticas até às 12 h do dia anterior

26

26

 **Leitura para a Próxima Aula**

BALANÇO DE ENERGIA EM ECOSISTEMA AMAZÔNICO POR MODELO DE REGRESSÃO ROBUSTA COM *BOOTSTRAP* E VALIDAÇÃO CRUZADA

Aguiar (2013)

Disponível na Página Pessoal – Publicações

Ler no mínimo p. 25, 26, 39 até 43 e 48 até 58

Estatística II

27

27

 **Listas 4 e 5**




Fazer parte no computador

Estatística II

28


28

 **Correlação e Regressão**

Estatística II

29

29

 **Correlação Linear**

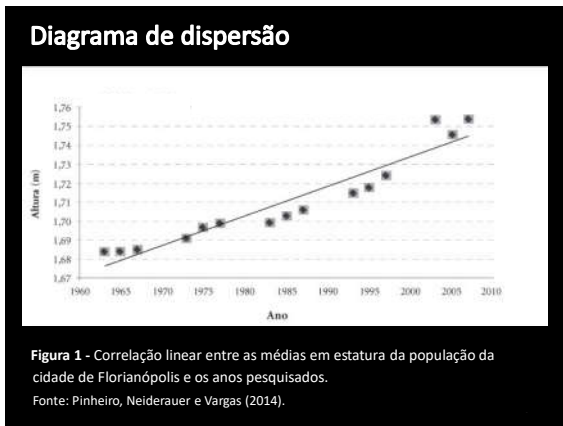
Para se medir o grau de correlação entre duas variáveis usa-se o coeficiente de correlação (r), que varia de -1 a $+1$.

Importante construir um diagrama de dispersão.

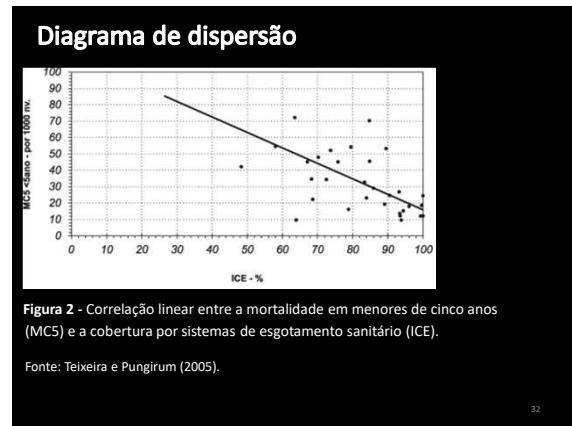
Estatística II

30

30



31



32

Correlação Linear

UNIR

Estatística II

Coefficiente de correlação de Pearson

Cuidado: causa-e-efeito

33

Correlação Linear

UNIR

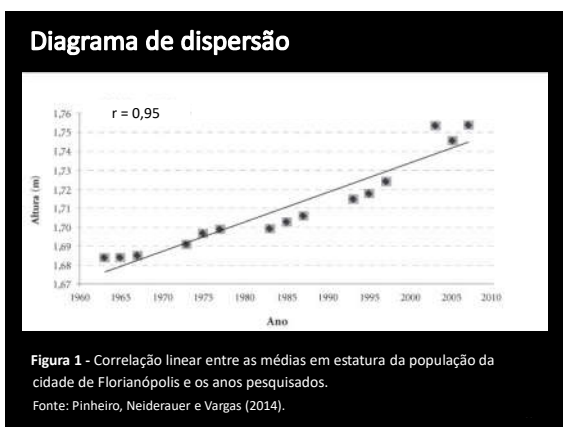
Estatística II

O coeficiente de correlação é dado por:

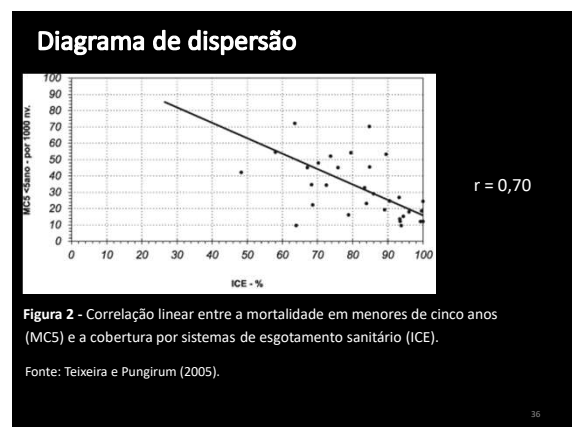
$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) \left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}}$$

n é o número de pares (x, y)

34



35



36

UNIR

Correlação Linear

Estatística II

Tabela 1 – Avaliação qualitativa de r quanto à intensidade

r	A correlação é dita
0	Nula
0,1	Fraca
0,3	Regular
0,6	Forte
0,9	Muito forte
1,0	Plena ou perfeita

37

37

UNIR

Correlação Linear

Estatística II

O que diferencia do que estudamos em Estatística I?

Testar a significância e verificar os pressupostos.

Veremos os pressupostos na aula no laboratório.

38

38

UNIR

Correlação Linear

Estatística II

Hipóteses

$H_0 : \rho = 0$

$H_1 : \rho \neq 0$

39

39

UNIR

Regressão Linear Simples

Estatística II

Causa

+

Efeito

→

Expressar
por fórmula

40

40

UNIR

Regressão Linear Simples

Estatística II

O estudo da regressão aplica-se àquelas situações em que há razões para supor uma relação de causa-efeito entre duas variáveis quantitativas e se deseja expressar matematicamente essa relação.

41

41

UNIR

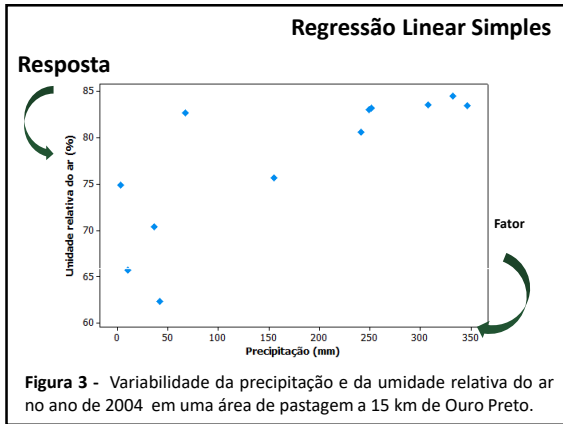
Regressão Linear Simples

Estatística II

Geralmente denomina-se a variável dependente (ou variável resposta) de y e a independente (fator) de x.

42

42



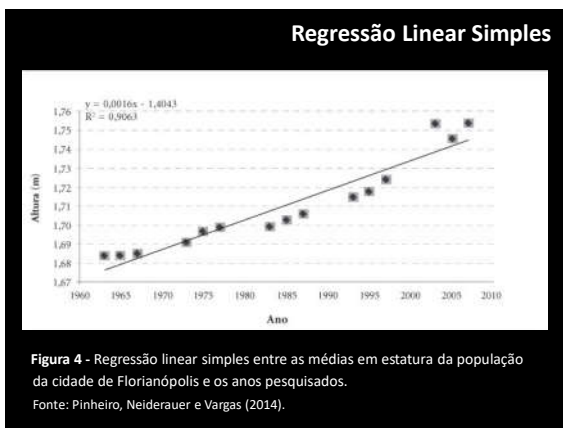
43

Regressão Linear Simples

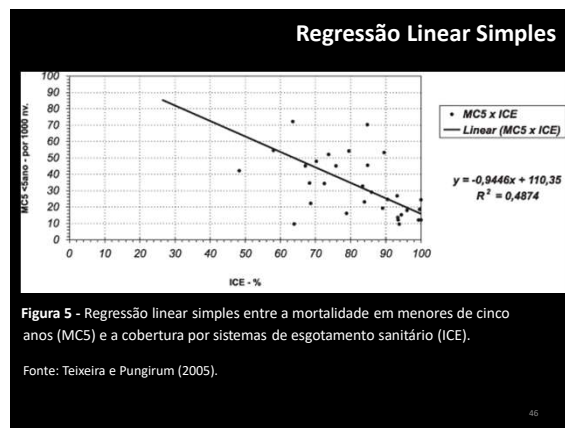
As expressões a seguir têm todas basicamente o mesmo significado:

- a) y depende de x (linguagem coloquial);
- b) y é função de x (linguagem matemática);
- c) existe regressão de y sobre x (linguagem estatística).

44



45



46

Regressão Linear Simples

Equação da reta

$Y = \alpha + \beta X + \epsilon$

Em que: Y é a variável dependente; α é o intercepto (coeficiente linear); β é uma medida da inclinação da reta (coeficiente de regressão ou angular); X é a variável independente do modelo e ϵ representa o erro aleatório.

47

47


Regressão Linear Simples

A reta estimada de regressão é:

$\hat{y} = a + bx$

48

48



Regressão Linear Simples


$$b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$a = \bar{y} - b\bar{x}$$

n é o número de pares (x, y)

49

49



Utilidades da Reta de Regressão

1. Permite representar a dependência de uma variável quantitativa em relação à outra por meio de uma equação simples.

50

50




Utilidades da Reta de Regressão

2. Permite prever valores para a variável dependente de acordo com valores determinados (inclusive não-observados) da variável independente.

Cuidado com a extrapolação!

51

51




Coefficiente de Determinação

Informa que fração da variabilidade de uma característica é explicada estatisticamente pela outra variável.

É expresso pelo quadrado do coeficiente de correlação, r^2 .

52

52



Regressão Linear Simples


O que diferencia do que estudamos em Estatística I?

Testar a significância e verificar os pressupostos.

Veremos os pressupostos na aula no laboratório.

53

53



Regressão Linear

Hipóteses

$$H_0 : \beta = 0$$

$$H_1 : \beta \neq 0$$

54

54

Situação-Problema 23

O departamento de saúde de uma grande cidade do Sudeste desenvolveu um índice de poluição do ar que mede o nível de vários poluentes do ar que causam doenças respiratórias nos seres humanos. A Tabela 2 fornece o índice de poluição (em uma escala de 1 a 10, onde 10 corresponde ao nível mais elevado de

55

Situação-Problema 23

poluentes) correspondente a sete dias do mês de agosto, selecionados aleatoriamente, bem como o número de pacientes com problema respiratório agudo que deram entrada na sala de emergência dos hospitais da cidade.

56

Situação-Problema 23

Tabela 2 – Índice de poluição do ar e o número de pacientes que deram entrada na emergência com problema respiratório agudo

Índice de poluição do ar	4,5	6,7	8,2	5,0	4,6	6,1	3,0
Atendimentos na emergência	53	82	102	60	39	42	27

a. Construa um diagrama de dispersão, encontre o coeficiente de correlação e teste se é significativo em um programa estatístico.

57

Situação-Problema 23

b. Desenvolva uma equação de regressão estimada para esses dados.

c. Encontre o coeficiente de determinação e verifique se é significativo.

58

Situação-Problema 23

d. Ache a melhor predição para o número de pacientes a dar entrada no hospital com problema respiratório agudo quando o índice de poluição do ar for 9.

e. Qual é a vantagem de ser capaz de determinar o número de pacientes a dar entrada no hospital com problema respiratório agudo a partir do índice de poluição do ar?


59

Situação-Problema 23

Um abraço fraterno e laranja ;)



60



Referências

Estatística II

AGUIAR, R. G. **Balço de Energia em Ecossistema Amaznico por Modelo de Regressão Robusta com Bootstrap e Validação Cruzada**. 85 f. Tese (Doutorado em Física Ambiental) – Instituto de Física, Universidade Federal de Mato Grosso, Cuiabá, 2013.

ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. **Estatística aplicada à Administração e Economia**. 2. ed. São Paulo: Pioneira Thomson Learning, 2003.

61

61



Referências

Estatística II


BECKER, J.; H. PABST, H.; MNYONGA, J.; KUZYAKOV, Y. Annual litterfall dynamics and nutrient deposition depending on elevation and land use at Mt. Kilimanjaro. **Biogeosciences**, v. 12, p. 5635-5646, 2015.

BUSSAB, W.O.; MORRETIN, P. A. **Estatística Básica**. São Paulo: Saraiva, 2003.

CALLEGARI-JACQUES, S. **Bioestatística: princípios e aplicações**. São Paulo: ARTMED, 2003.

62

62



Referências

Estatística II

COSTA, S. F. **Introdução ilustrada à Estatística**. 4. ed. São Paulo: Harbra, 2005.


CRESPO, A. A. **Estatística fácil**. 17. ed. São Paulo: Saraiva, 1999.

FREUND, J. E.; SIMON, G. A. **Estatística aplicada: Economia, Administração e Contabilidade**. 9. ed. Porto Alegre: Bookman, 2000.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 5. ed. São Paulo: Atlas, 2007.

63

63



Referências


Estatística II

LEVIN, J.; FOX, J. A. **Estatística para ciências humanas**. 9. ed. São Paulo: Prentice Hall, 2004.

MONTGOMERY, D. C.; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros**. 4. ed. Rio de Janeiro: LTC, 2009.

64

64



Referências


Estatística II

PINHEIRO, A. C. C.; NIEDERAUER, J. M.; VARGAS, D. M. Tendência secular de crescimento em estatura no município de Florianópolis (SC), Brasil, e sua associação com o índice de desenvolvimento humano (IDH). **Ciência & Saúde Coletiva**, Rio de Janeiro, v. 19, n. 1, p. 227234, jan. 2014. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S141381232014000100227&lng=en&nrm=iso>. Acesso em: 29 out. 2019. <http://dx.doi.org/10.1590/1413-81232014191.1913>.

SPIEGEL, M. R. **Estatística: resumo da teoria, 975 problemas resolvidos, 619 problemas propostos**. São Paulo: McGraw-Hill do Brasil, 1975.

65

65




Referências

Estatística II

TEIXEIRA, J. C.; PUNGIRUM, M. E. M. C. Análise da associação entre saneamento e saúde nos países da América Latina e do Caribe, empregando dados secundários do banco de dados da Organização Pan-Americana de Saúde - OPAS. **Revista Brasileira de Epidemiologia**, São Paulo, v. 8, n. 4, p. 365-376, dez. 2005. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-790X2005000400005&lng=en&nrm=iso>. Acesso em: 29 out. 2019. <http://dx.doi.org/10.1590/S1415-790X2005000400005>.

66

66



Referências

TRIOLA, M. F. **Introdução à Estatística**. 10. ed. Rio de Janeiro: LTC, 2008.

VIEIRA, S. **Análise de Variância (ANOVA)**. São Paulo: Atlas, 2006.

67

Estadística II

67