



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



Estatística I


Profa. Renata Gonçalves Aguiar



Despertando o (a) Discente Ativo(a)




Distribuição Normal



Distribuição Normal

É um clássico da Estatística

5



Distribuição Normal

Qual o tipo de distribuição de probabilidade dos dados?

Quais os parâmetros?

6

Distribuição Normal

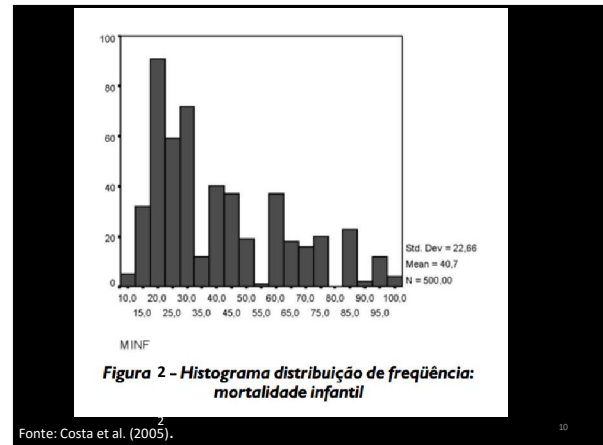
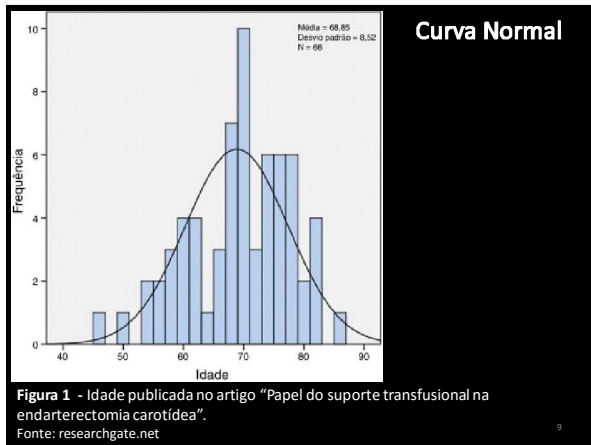
Esta é a mais importante distribuição de probabilidade para descrever uma variável aleatória contínua.

7

Distribuição Normal

A curva normal tem dois parâmetros, μ e σ , $N(\mu, \sigma)$. Eles determinam a posição e a forma da distribuição.

8

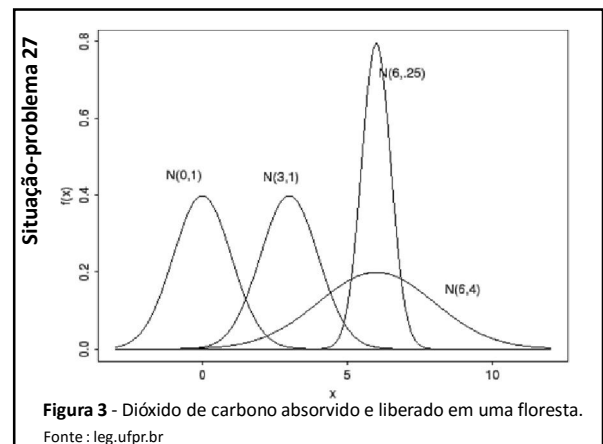


Situação-problema 27

Suponha que sejam realizadas quatro coletas sobre o CO₂ absorvido e liberado por uma área de floresta (t C ha⁻¹ a⁻¹) e os resultados se apresentem de acordo com a Figura 3.

Quais conclusões podemos chegar?

11



Função de Densidade Normal de Probabilidade

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\pi = 3,14159$$

$$e = 2,71828$$

13

Características Gerais

1. A variável aleatória pode assumir qualquer valor real.
2. O gráfico da distribuição normal é uma curva em forma de sino, simétrica em torno da média.

14

Características Gerais

3. A área total sob a curva é 1.
4. Como a curva é simétrica em torno da média, os valores maiores e menores que a média ocorrem com igual probabilidade.

15

Características Gerais

5. O ponto mais alto na curva está na média, que também é a mediana e a moda da distribuição.
6. O desvio padrão determina a largura da curva.

16

Características Gerais

7. As probabilidades para a variável aleatória normal são dadas por áreas sob a curva. A porcentagem de valores em alguns intervalos comumente usados são:

- a) 68,26% dos valores de uma variável aleatória normal estão dentro de um desvio padrão positivo e negativo de sua média.

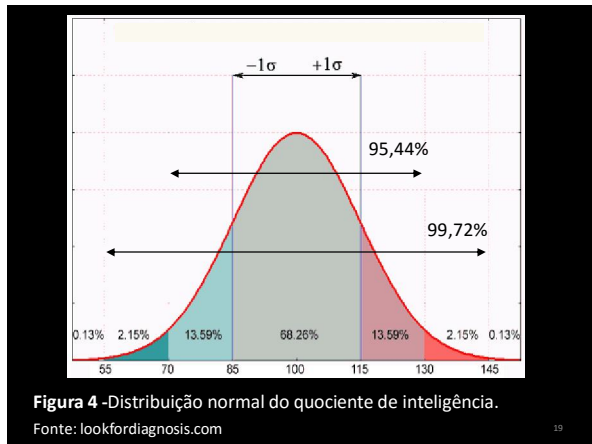
17

Características Gerais

b) 95,44% dos valores de uma variável aleatória normal estão dentro de dois desvios padrões positivos e negativos de sua média.

c) 99,72% dos valores de uma variável aleatória normal estão dentro de três desvios padrões positivos e negativos de sua média.

18



Situação-problema 28

A concentração de carbono orgânico total (COT) na água apresenta valor alerta (valor acima do qual representa possibilidade de causar prejuízos ao ambiente na área de disposição) de 10% de acordo com a Resolução Conama n. 344, de 25 de março de 2004.

Situação-problema 28

Suponha que o rio Madeira tenha distribuição $N(8; 1,5)$. Qual a chance, de que em um dado dia, a concentração de COT exceda o valor alerta?

Situação-problema 28

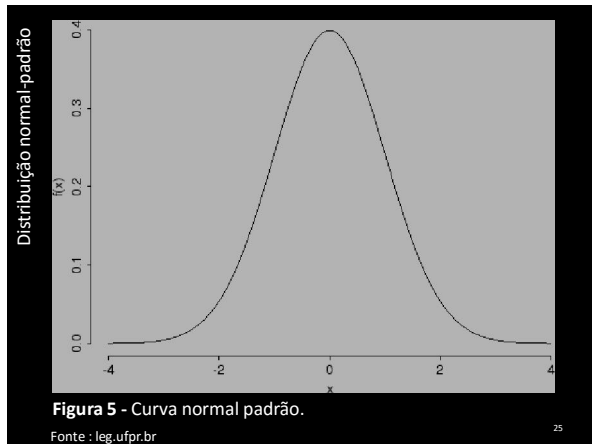
Aplicando a equação...

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Distribuição Normal-Padrão de Probabilidade

Distribuição Normal-Padrão

Denomina-se distribuição normal-padrão a distribuição normal de média zero e desvio padrão de 1.



Cálculo das Probabilidades

Fórmulas usadas para converter qualquer variável aleatória normal para a distribuição normal:

$$z = \frac{x - \mu}{\sigma} \qquad z = \frac{x - \bar{x}}{s}$$

Retomando a S-P 28

Suponha que o rio Madeira tenha distribuição $N(8; 1,5)$. Qual a chance, de que em um dado dia, a concentração de COT exceda o valor alerta?

Situação-Problema 29

Suponha que as notas de um vestibular tenham distribuição normal com média 60 e desvio padrão de 15 pontos.

a. Se você prestou esse vestibular e obteve 80 pontos, qual é a sua posição relativa, em unidades de desvio padrão, com relação à média de notas?

Situação-Problema 29

- b. Qual a probabilidade de um candidato ter obtido nota acima de 55 pontos?
- c. Se foram considerados aprovados os candidatos que obtiveram nota mínima correspondente a um desvio padrão acima da média, qual é a nota mínima de aprovação na escala original dos dados?

Situação-Problema 30

Dado que z é uma variável aleatória normal-padrão, calcule as seguintes probabilidades.

- a. $P(0 \leq z \leq 1)$.
- b. $P(-1 \leq z \leq 1)$.
- c. de ocorrer valor menor do que $z = -0,50$.

Artigo para a aula de hoje

Tratamento de águas residuárias de suinocultura em sistemas alagados construídos, com *Chrysopogon zizanioides* e *Polygonum punctatum* cultivadas em leito de argila expandida

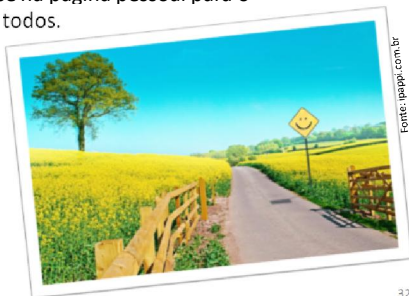
*Swine wastewater treatment in constructed wetlands cultivated with *Chrysopogon zizanioides* and *Polygonum punctatum* on expanded clay bed*

Nilton de Freitas Souza Ramos¹, Allisson Carraro Borges², Gustavo Castro Gonçalves³, Antonio Teixeira de Matos⁴

31

Lista 4

Encontra-se na página pessoal para o leite de todos.



32

Correlação e Regressão

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.

34

Diagrama de dispersão

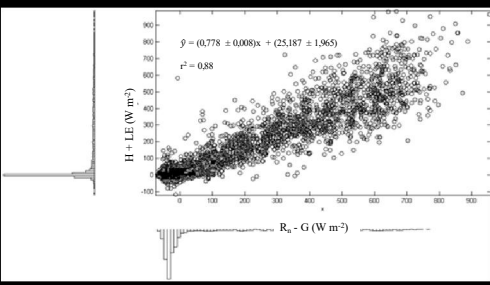


Figura 6 - Relação entre a energia disponível no sistema ($R_n - G$) e a soma dos fluxos de calor sensível e calor latente ($H + LE$). R_n é o saldo de radiação e G é o fluxo de calor no solo. Os pontos representam médias a cada 30 min, $n = 5.320$.
Fonte: Aguiar (2013).

35

Correlação Linear

Cuidado: causa-e-efeito.

36

Correlação Linear

Coeficiente de correlação de Pearson

37

Correlação Linear

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)

38

Correlação Linear

Tabela 1 – Avaliação qualitativa de r quanto à intensidade

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

39

Situação-Problema 31

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

40

Situação-Problema 31

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

41

Situação-Problema 31

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

Construa um diagrama de dispersão e encontre o coeficiente de correlação.

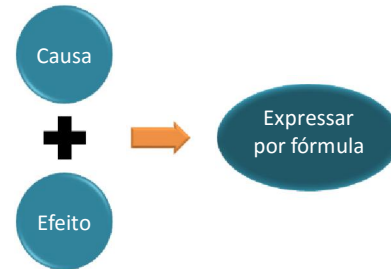
42

Regressão Linear Simples

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.

43

Regressão Linear Simples



44

Regressão Linear Simples

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

45

Regressão Linear Simples

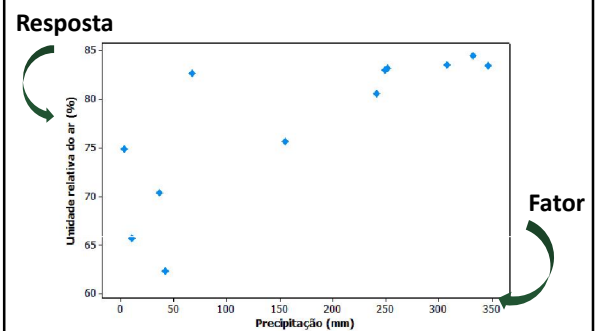


Figura 7 - Variabilidade da precipitação e da umidade relativa do ar no ano de 2004 em uma área de pastagem a 15 km de Ouro Preto.

Regressão Linear Simples

As expressões a seguir têm todas basicamente o mesmo significado:

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

47

Regressão Linear Simples



48

Regressão Linear Simples

A equação da reta é dada por

$$Y = A + Bx.$$

Assim, a reta estimada de regressão é:

$$\hat{y} = a + bx$$

49

Regressão Linear Simples

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

n é o número de pares (x, y)

50

Utilidades da Reta de Regressão

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

51

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!

52

Coeficiente 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 .

53

Situação-Problema 32

Com base nos dados da atividade 31, faça o que se pede.

- a. Desenvolva uma equação de regressão estimada para esses dados.
- b. Encontre o coeficiente de determinação.

54

Situação-Problema 32

c. 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. 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?

55

Importantíssimo

Para realizar uma correlação e/ou uma regressão precisam ser atendidos alguns pressupostos que serão estudados em Estatística II.

56

Referências

AGUIAR, R. G. **Balanco de Energia em Ecossistema Amazônico 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.

BARBETTA, P. A. **Estatística aplicada às Ciências Sociais**. 5. ed. Florianópolis: Ed. da UFSC, 2002.

57

Referências

BUSSAB, W. O.; MORRETIN, P. A. **Estatística Básica**. 5. ed. São Paulo: Saraiva, 2003.

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

COSTA, S. S.; HELLER, L.; BRANDAO, C. C. S.; COLOSIMO, E. A. Indicadores epidemiológicos aplicáveis a estudos sobre a associação entre saneamento e saúde de base municipal. **Engenharia Sanitária e Ambiental** [online], Rio de Janeiro, v. 10, n. 2, p. 118-127, abr./jun. 2005.

58

Referências

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.

59

Referências

MONTGOMERY, D. C.; RUNGER, G. C. **Estatística Aplicada e Probabilidade para Engenheiros**. 4. ed. Rio de Janeiro: LTC, 2009.

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

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

60