**Universidade Federal de Pernambuco**

**Centro de Informática**

**Curso de Graduação em Ciência da Computação**

**Estatística p/ Computação**

**Professora: Renata**

**Projeto de Estatística: Abalone**

**Bruno Machado Dias Macena (bmdm)**

**Igor Gomes (ilrg)**

**Jonathas Mendes Alves (jma)**

**Renato Oliveira (rso)**

**Recife, novembro de 2009**

**Bruno Machado Dias Macena (bmdm)**

**Igor Gomes (ilrl)**

**Jonathas Alves (jma)**

**Renato Oliveira (rso)**

**Projeto de Estatística: Abalone**

Trabalho complementar ao seminário que serve como pré-requisito para obtenção da 3ª nota na disciplina de Estatística p/ Computação no curso de Ciência da Computação, sob a orientação da professora Renata Souza.

**Recife, novembro de 2009**

**Sumário**

[1 Introdução 4](#_Toc245971103)

[2 Desenvolvimento 5](#_Toc245971104)

[3 Referências 6](#_Toc245971105)

# Introdução

 O seguinte relatório tem como objetivo estudar formas alternativas de detecção da idade dos Abalones, baseando-se em suas medidas.

 [*Haliotis*](http://pt.wikipedia.org/wiki/Haliotis) ou abalone é um [gastrópode](http://pt.wikipedia.org/wiki/Gastr%C3%B3pode), o único na família [Haliotidae](http://pt.wikipedia.org/w/index.php?title=Haliotidae&action=edit&redlink=1). Sua carne é um prato muito apreciado na [Ásia](http://pt.wikipedia.org/wiki/%C3%81sia).

 Atualmente, a forma usada para estimar sua idade consiste em fazer um corte transversal da sua casca através do seu cone. A amostra é então colorida e analisada em um microscópio. A idade é estimada contando o número de anéis da amostra. Isso é uma tarefa longa e tediosa.

 O novo método proposto para estimação da idade do abalone é baseado em outras medidas que são mais fáceis de obter, tais como altura, diâmetro, largura dos orifícios, sexo e anéis. Outras informações como padrões climáticos e localização também podem ser usadas para a estimação.

 Para este relatório, foi escolhida a variável discreta número de anéis e as variáveis contínuas: tamanho, diâmetro, altura e comprimento da crosta.

 Foram medidas moda, mediana, média, variância, desvio padrão, coeficiente de assimetria, coeficiente de curtose, primeiro quartil e terceiro quartil.

# Desenvolvimento

Como já antecipado na introdução, as variáveis escolhidas para análise foram a discreta número de anéis e as contínuas : tamanho, diâmetro, altura e comprimento da crosta.

O motivo para a escolha destas foi que elas são mais significativas na proposta do trabalho, já que sexo, tamanho dos orifícios e tamanho das vísceras não são determinantes para a estimativa da idade.

A seguir cada variável será analisada estatisticamente.

## Comprimento

|  |  |
| --- | --- |
| MEDIDA ESTATÍSTICA | VALOR |
| MÉDIA | 0.524 |
| MODA |  |
| MEDIANA | 0.545 |
| VARIÂNCIA | 0.01442231 |
| DESVIO PADRÃO | 0.1200929 |
| COEF. DE ASSIMETRIA |  |
| COEF. DE CURTOSE | 3.061641 |
| 1º QUARTIL | 0.450 |
| 3º QUARTIL | 0.615 |

|  |  |
| --- | --- |
| MEDIDA ESTATÍSTICA | VALOR |
| MÉDIA | 0.4079 |
| MODA |  |
| MEDIANA | 0.4250 |
| VARIÂNCIA | 0.009848551 |
| DESVIO PADRÃO | 0.09923987 |
| COEF. DE ASSIMETRIA |  |
| COEF. DE CURTOSE | 2.951729 |
| 1º QUARTIL | 0.35 |
| 3º QUARTIL | 0.48 |

## Diâmetro

## Altura

|  |  |
| --- | --- |
| MEDIDA ESTATÍSTICA | VALOR |
| MÉDIA | 0.1395 |
| MODA |  |
| MEDIANA | 0.14 |
| VARIÂNCIA | 0.001749503 |
| DESVIO PADRÃO | 0.04182706 |
| COEF. DE ASSIMETRIA |  |
| COEF. DE CURTOSE | 78.89531 |
| 1º QUARTIL | 0.115 |
| 3º QUARTIL | 0.165 |

## Peso da casca

|  |  |
| --- | --- |
| MEDIDA ESTATÍSTICA | VALOR |
| MÉDIA | 0.2388 |
| MODA |  |
| MEDIANA | 0.234 |
| VARIÂNCIA | 0.01937738 |
| DESVIO PADRÃO | 0.1392027 |
| COEF. DE ASSIMETRIA |  |
| COEF. DE CURTOSE | 3.528164 |
| 1º QUARTIL | 0.13 |
| 3º QUARTIL | 0.329 |

## Anéis

|  |  |
| --- | --- |
| MEDIDA ESTATÍSTICA | VALOR |
| MÉDIA | 9.934 |
| MODA |  |
| MEDIANA | 9 |
| VARIÂNCIA | 10.39527 |
| DESVIO PADRÃO | 3.224169 |
| COEF. DE ASSIMETRIA |  |
| COEF. DE CURTOSE | 5.323912 |
| 1º QUARTIL | 8 |
| 3º QUARTIL | 11 |

## Tamanho da amostra: **4177**

## **Graficos e tabelas:**

##  **Histograma do Comprimento**

##

## **Histograma do Diâmetro**

##

## **Histograma da Altura:**

## **FALTANDO (hist(abalone$Altura))**

## **Histograma da Largura da casca (eh PESO DA CASCA):**

##

## **Histograma dos Aneis:**

##

## **Frequencia**

## **Comprimento**

## FALTANDOALTURA

## Peso

## Diametro

##

## Peso da casca

##

## Aneis

##

## *Obs. Novos métodos de calcular o número de classes do histograma devem ser estudados e tais histogramas devem ser comparados em cada variável (ambos histogramas deverão estar no relatório). Além disso, as tabelas de freqüência podem ser feita no Excel, no entanto todos os valores contidos na tabela deverão ser calculados em R.*

## **QUAIS SÃO ESSES NOVOS METODOS??**

## **FALTA FAZER A TABELA DE FREQUENCIA!!**

# Referências

MEIDANIS, João. **Bioinformática na indústria.** Disponível em < http://www.comciencia.br/reportagens/bioinformatica/bio14.shtml>. Acesso em: 02 de jun. de 2009.

ALECRIM, Emerson. **Redes Neurais Artificiais.** Disponível em < http://www.infowester.com/redesneurais.php>. Acesso em: 03 de jun. de 2009.

TATIBANA, Cássia; KAETSU, Deisi. **Uma introdução às Redes Neurais.** Disponível em < http://www.din.uem.br/ia/neurais/#utiliza%C3%A7%C3%A3o>. Acesso em: 03 de jun. de 2009.

EMBRAPA. **Sobre o LBA / Projetos em Andamento.** Disponível em < http://www.lba.cnptia.embrapa.br/lba/index.html>. Acesso em: 03 de jun. de 2009.

**LINHA DE PESQUISA: BIOINFORMÁTICA COMPLEXA.** Disponível em < http://www.bioinfo.ufpr.br/mod/resource/view.php?id=37>. Acesso em: 02 de jun. de 2009.

Grupo de Ciências Biológicas do IST. **PCR – Amplificação de DNA in vitro.** Disponível em <http://www.e-escola.pt/topico.asp?hid=339>. Acesso em: 01 de jun. de 2009.

**Eletroforese.** Disponível em <http://www.icb.ufmg.br/mor/pad-morf/eletroforese.htm>. Acesso em: 01 de jun. de 2009.

Applied BioSystems. **The Basics: Northern Analysis.** Disponível em < http://www.ambion.com/techlib/basics/northerns/index.html>. Acesso em: 01 de jun. de 2009.

Sítio Centro de Informática (CIn). **Perfil - Bioinformática**. Disponível em < http://www.cin.ufpe.br/~graduacao/reforma/a\_perfil\_bio-informatica.htm>. Acesso em: 04 de jun. de 2009.

WIECZOREK, Emilio; LEAL, Eduardo. **Caminhos e Tendências do uso de Banco de Dados na Bioinformática.** Disponível em <http://www.wieczorek.com.br/publicacoes/artigo\_IVencoinfo.pdf>. Acesso em: 04 de jun. de 2009.

SETUBAL, João. **A origem e o sentido da bioinformática.** Disponível em <http://www.comciencia.br/reportagens/bioinformatica/bio10.shtml>. Acesso em: 02 de jun. de 2009.

OTTO, Thomas. **A plataforma PDTIS de bioinformática.** Disponível em <http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/98/107>. Acesso em: 16 de jun. de 2009.

MIRANDA, Márcio. **Algoritmos genéticos: Fund. e aplicações.** Disponível em < http://www.gta.ufrj.br/~marcio/genetic.html>. Acesso em: 03 de jun. de 2009.

OBITKO, Marek. **ALGORITMOS GENÉTICOS.** Disponível em <http://professor.webizu.org/ga>. Acesso em: 17 de jun. de 2009.

PERONE, Christian. **Pyevolve Documentation.** Disponível em <http://pyevolve.sourceforge.net>. Acesso em: 16 de jun. de 2009.