Relatório de Análise da Assessoria (conclusão)

Porto Alegre, 14 de outubro de 2020

# Apresentação

Este documento tem como objetivo apresentar os métodos estatísticos para análise de dados do trabalho de doutorado **“Spatiotemporal patterns of carrion biomass of marine tetrapods in the interface ocean-land in the southern Brazilian coastline”** do assessorado **Maurício Tavares** do **Programa de Pós-Graduação em Biologia Animal** do **Instituto de Biociências** da **Universidade Federal do Rio Grande do Sul**.

A seguir são apresentados os objetivos do trabalho de doutorado, as variáveis do estudo e análise estatística.

## Equipe

**Professor:** Rodrigo Citton Padilha dos Reis

**Alunos:**

* Celso Menoti da Silva
* Enzo Bertoldi Oestreich
* Nicolas Mathias Hahn
* Victor Frank Gomes

# Introdução

## Objetivos

### Objetivo geral

Fornecer uma quantificação robusta e abrangente da biomassa de carcaças no importante e ainda pouco estudado ecossistema: a interface oceano-terra.

### Objetivos específicos

* Descrever a distribuição espaço-temporal da biomassa de carcaças de animais marinhos encalhados na costa litorânea brasileira.
* Estimar a quantidade de biomassa das carcaças dos animais marinhos encalhados (separada por espécie e multiplicada pela abundância de cada uma).

# Análise estatística

A análise estatísica empregada no artigo *“Spatiotemporal patterns of carrion biomass of marine tetrapods in the interface ocean-land in the southern Brazilian coastline”* responde aos objetivos do mesmo.

## Sugestões de análise

* Verificação do código da análise estatística;
* Reanálise a partir do código da análise estatística;
* **Análise de dados circulares:** descrição dos dados anuais de abundância e riqueza de espécies, pode ser realizado teste de distribuição uniforme ao longo do ano (contra a alternativa de agrupamentos em torno de alguns períodos específicos no tempo).
  + Teste de Rayleigh e de Watson;
  + Pacote em R circular
* Estimativa do erro padrão relativo.

# Análises

#--------------------#  
# Carregando Pacotes #  
#--------------------#  
  
library(circular)  
  
#----------------#  
# Banco de Dados #  
#----------------#  
  
data <- read.csv("dados mortalidade.csv",h=T,sep=",")  
head(data)

## group angle abundance biomass  
## 1 Mysticeti 2 2 9975.0924  
## 2 Mysticeti 11 1 14081.7648  
## 3 Mysticeti 18 1 6255.1603  
## 4 Mysticeti 26 1 3719.9322  
## 5 Mysticeti 30 1 811.1605  
## 6 Mysticeti 43 1 3719.9322

data$group = as.factor(data$group) # transformando em fator  
taxa <- levels(data[,1])  
  
# Grupos  
Mysticeti <- data[data$group=="Mysticeti",]  
Odontoceti <- data[data$group=="Odontoceti",]  
Pinipedia <- data[data$group=="Pinipedia",]   
Procellariiformes <- data[data$group=="Procellariiformes",]  
SeaTurtles <- data[data$group=="Sea Turtles",]  
Sphenisciformes <- data[data$group=="Sphenisciformes",]

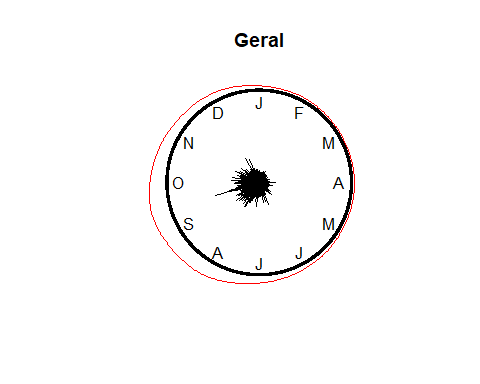
## Código da Análise Estatística

Obs.: Realizaremos comentários sobre o código ou algo assim? Questiono porque me parece que nenhuma premissa estatística foi quebrada ao observar o código, e utilizei o código deles como base para montar o de “Estatística Circular”

## Estatística Circular

### Grupo Geral

geral <- rep(data$angle, data$abundance)  
geral\_rad <- rad(geral)  
geral\_circ <- as.circular(geral\_rad)  
  
## Gráficos  
  
# frequência  
rose.diag(geral\_circ, rotation = "clock", zero = pi/2,units = "rads",   
 axes = F, border = "black", ticks = T, prop = 3, main = "Geral")  
axis.circular(at=circular(sort(seq(0, 11/6\*pi, pi/6), decreasing = T)), c(labels = c("M", "J",  
 "J", "A", "S", "O", "N", "D", "J", "F", "M","A")))  
lines(density(geral\_circ, bw = 20), col = "red", rotation = "clock", zero = pi/2, shrink = 1.75)



Como pode-se notar pelo gráfico acima, temos uma maior concentração no período de Agosto(A) a Novembro(N). Podemos realizar teste de hipóteses para averiguar se, de fato, há uma diferença na distribuição dos dados ao longo do ano.

## Testes de Hipóteses  
  
rao.spacing.test(geral\_circ)

##   
## Rao's Spacing Test of Uniformity   
##   
## Test Statistic = 357.5613   
## P-value < 0.001   
##

rayleigh.test(geral\_circ)

##   
## Rayleigh Test of Uniformity   
## General Unimodal Alternative   
##   
## Test Statistic: 0.462   
## P-value: 0

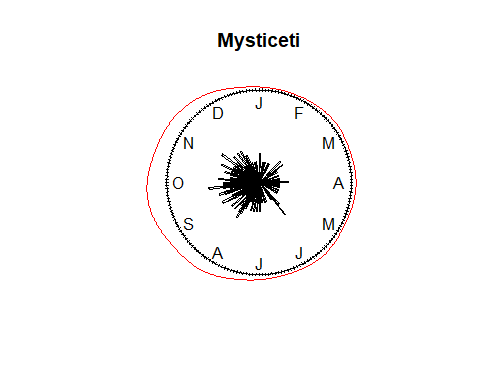
watson.test(geral\_circ, dist = "uniform")

##   
## Watson's Test for Circular Uniformity   
##   
## Test Statistic: 587.696   
## P-value < 0.01   
##

Como podemos notar pelos testes acima, rejeitamos a hipótese de os dados estarem distribuídos uniformemente ao longo do ano.

### Mysticeti

mysti <- rep(Mysticeti$angle, Mysticeti$abundance)  
mysti\_rad<-rad(mysti)  
mysti\_circ<-as.circular(mysti\_rad)  
  
  
## Gráficos  
  
# frequência  
rose.diag(mysti\_circ, rotation = "clock", zero = pi/2,units = "rads",   
 axes = F, border = "black", ticks = T, prop = 3, main = "Mysticeti")  
axis.circular(at=circular(sort(seq(0, 11/6\*pi, pi/6), decreasing = T)), c(labels = c("M", "J",  
 "J", "A", "S", "O", "N", "D", "J", "F", "M","A")))  
lines(density(mysti\_circ, bw = 20), col = "red", rotation = "clock", zero = pi/2, shrink = 1.75)



Como pode-se notar pelo gráfico acima, temos uma maior concentração no período de Agosto(A) a Novembro(N). Podemos realizar teste de hipóteses para averiguar se, de fato, há uma diferença na distribuição dos dados ao longo do ano.

## Testes de Hipóteses  
  
rao.spacing.test(mysti\_circ)

##   
## Rao's Spacing Test of Uniformity   
##   
## Test Statistic = 166   
## P-value < 0.001   
##

rayleigh.test(mysti\_circ)

##   
## Rayleigh Test of Uniformity   
## General Unimodal Alternative   
##   
## Test Statistic: 0.4359   
## P-value: 0

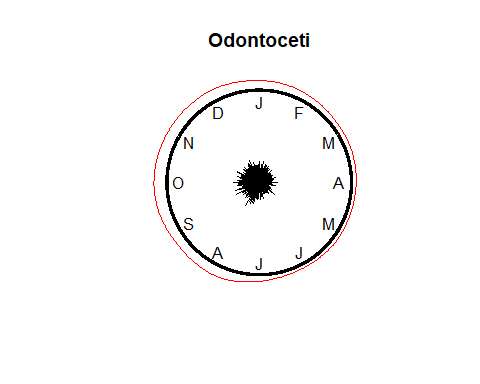
watson.test(mysti\_circ, dist = "uniform")

##   
## Watson's Test for Circular Uniformity   
##   
## Test Statistic: 1.8779   
## P-value < 0.01   
##

Como podemos notar pelos testes acima, rejeitamos a hipótese de os dados estarem distribuídos uniformemente ao longo do ano.

### Odontoceti

odonto <- rep(Odontoceti$angle, Odontoceti$abundance)  
odonto\_rad<-rad(odonto)  
odonto\_circ<-as.circular(odonto\_rad)  
  
## Gráficos  
  
# frequência  
rose.diag(odonto\_circ, rotation = "clock", zero = pi/2,units = "rads",   
 axes = F, border = "black", ticks = T, prop = 3, main = "Odontoceti")  
axis.circular(at=circular(sort(seq(0, 11/6\*pi, pi/6), decreasing = T)), c(labels = c("M", "J",  
 "J", "A", "S", "O", "N", "D", "J", "F", "M","A")))  
lines(density(odonto\_circ, bw = 20), col = "red", rotation = "clock", zero = pi/2, shrink = 1.75)



Como pode-se notar pelo gráfico acima, temos uma maior concentração no período de Agosto(A) a Dezembro(D). Podemos realizar teste de hipóteses para averiguar se, de fato, há uma diferença na distribuição dos dados ao longo do ano.

## Testes de Hipóteses  
  
rao.spacing.test(odonto\_circ)

##   
## Rao's Spacing Test of Uniformity   
##   
## Test Statistic = 319.0883   
## P-value < 0.001   
##

rayleigh.test(odonto\_circ)

##   
## Rayleigh Test of Uniformity   
## General Unimodal Alternative   
##   
## Test Statistic: 0.2645   
## P-value: 0

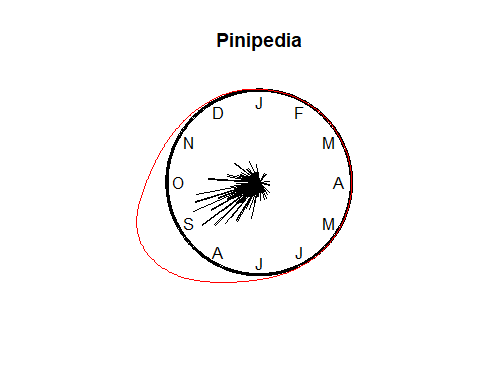
watson.test(odonto\_circ, dist = "uniform")

##   
## Watson's Test for Circular Uniformity   
##   
## Test Statistic: 11.1617   
## P-value < 0.01   
##

Como podemos notar pelos testes acima, rejeitamos a hipótese de os dados estarem distribuídos uniformemente ao longo do ano.

### Pinipedia

pini <- rep(Pinipedia$angle, Pinipedia$abundance)  
pini\_rad<-rad(pini)  
pini\_circ<-as.circular(pini\_rad)  
  
  
## Gráficos  
  
# frequência  
rose.diag(pini\_circ, rotation = "clock", zero = pi/2,units = "rads",   
 axes = F, border = "black", ticks = T, prop = 3, main = "Pinipedia")  
axis.circular(at=circular(sort(seq(0, 11/6\*pi, pi/6), decreasing = T)), c(labels = c("M", "J",  
 "J", "A", "S", "O", "N", "D", "J", "F", "M","A")))  
lines(density(pini\_circ, bw = 20), col = "red", rotation = "clock", zero = pi/2, shrink = 1.75)



Como pode-se notar pelo gráfico acima, temos uma maior concentração no período de Agosto(A) a Outubro(O). Podemos realizar teste de hipóteses para averiguar se, de fato, há uma diferença na distribuição dos dados ao longo do ano.

## Testes de Hipóteses  
  
rao.spacing.test(pini\_circ)

##   
## Rao's Spacing Test of Uniformity   
##   
## Test Statistic = 293.8776   
## P-value < 0.001   
##

rayleigh.test(pini\_circ)

##   
## Rayleigh Test of Uniformity   
## General Unimodal Alternative   
##   
## Test Statistic: 0.8342   
## P-value: 0

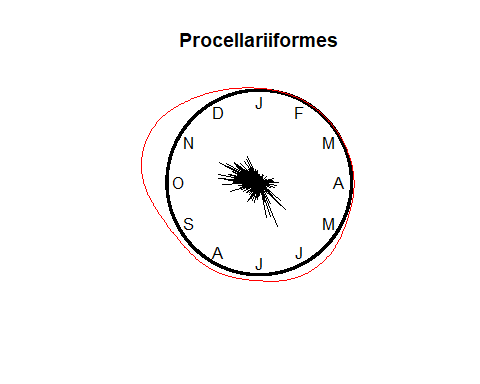
watson.test(pini\_circ, dist = "uniform")

##   
## Watson's Test for Circular Uniformity   
##   
## Test Statistic: 29.7304   
## P-value < 0.01   
##

Como podemos notar pelos testes acima, rejeitamos a hipótese de os dados estarem distribuídos uniformemente ao longo do ano.

### Procellariiformes

proce <- rep(Procellariiformes$angle, Procellariiformes$abundance)  
proce\_rad<-rad(proce)  
proce\_circ<-as.circular(proce\_rad)  
  
## Gráficos  
  
# frequência  
rose.diag(proce\_circ, rotation = "clock", zero = pi/2,units = "rads",   
 axes = F, border = "black", ticks = T, prop = 3, main = "Procellariiformes")  
axis.circular(at=circular(sort(seq(0, 11/6\*pi, pi/6), decreasing = T)), c(labels = c("M", "J",  
 "J", "A", "S", "O", "N", "D", "J", "F", "M","A")))  
lines(density(proce\_circ, bw = 20), col = "red", rotation = "clock", zero = pi/2, shrink = 1.75)



Como pode-se notar pelo gráfico acima, temos uma maior concentração no período de Outubro(O) a Dezembro(D). Podemos realizar teste de hipóteses para averiguar se, de fato, há uma diferença na distribuição dos dados ao longo do ano.

## Testes de Hipóteses  
  
rao.spacing.test(proce\_circ)

##   
## Rao's Spacing Test of Uniformity   
##   
## Test Statistic = 344.8542   
## P-value < 0.001   
##

rayleigh.test(proce\_circ)

##   
## Rayleigh Test of Uniformity   
## General Unimodal Alternative   
##   
## Test Statistic: 0.484   
## P-value: 0

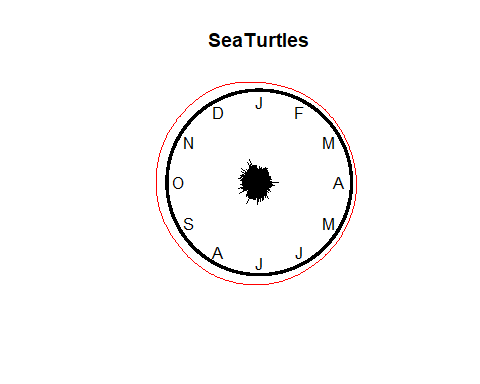
watson.test(proce\_circ, dist = "uniform")

##   
## Watson's Test for Circular Uniformity   
##   
## Test Statistic: 110.9215   
## P-value < 0.01   
##

Como podemos notar pelos testes acima, rejeitamos a hipótese de os dados estarem distribuídos uniformemente ao longo do ano.

### SeaTurtles

turtles <- rep(SeaTurtles$angle, SeaTurtles$abundance)  
turtles\_rad<-rad(turtles)  
turtles\_circ<-as.circular(turtles\_rad)  
  
## Gráficos  
  
# frequência  
rose.diag(turtles\_circ, rotation = "clock", zero = pi/2,units = "rads",   
 axes = F, border = "black", ticks = T, prop = 3, main = "SeaTurtles")  
axis.circular(at=circular(sort(seq(0, 11/6\*pi, pi/6), decreasing = T)), c(labels = c("M", "J",  
 "J", "A", "S", "O", "N", "D", "J", "F", "M","A")))  
lines(density(turtles\_circ, bw = 20), col = "red", rotation = "clock", zero = pi/2, shrink = 1.75)



Como pode-se notar pelo gráfico acima, temos uma maior concentração no período de Julho(J) a Janeiro(J). Podemos realizar teste de hipóteses para averiguar se, de fato, há uma diferença na distribuição dos dados ao longo do ano.

## Testes de Hipóteses  
  
rao.spacing.test(turtles\_circ)

##   
## Rao's Spacing Test of Uniformity   
##   
## Test Statistic = 354.4092   
## P-value < 0.001   
##

rayleigh.test(turtles\_circ)

##   
## Rayleigh Test of Uniformity   
## General Unimodal Alternative   
##   
## Test Statistic: 0.206   
## P-value: 0

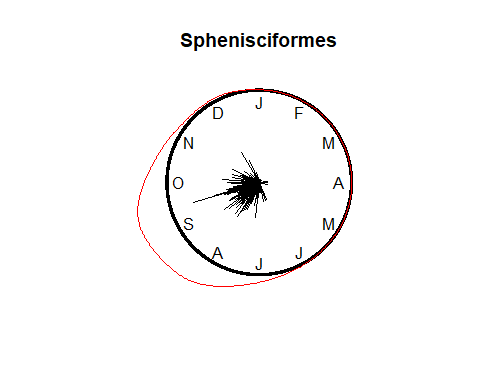
watson.test(turtles\_circ, dist = "uniform")

##   
## Watson's Test for Circular Uniformity   
##   
## Test Statistic: 51.0521   
## P-value < 0.01   
##

Como podemos notar pelos testes acima, rejeitamos a hipótese de os dados estarem distribuídos uniformemente ao longo do ano.

### Sphenisciformes

spheni <- rep(Sphenisciformes$angle, Sphenisciformes$abundance)  
spheni\_rad<-rad(spheni)  
spheni\_circ<-as.circular(spheni\_rad)  
  
  
## Gráficos  
  
# frequência  
rose.diag(spheni\_circ, rotation = "clock", zero = pi/2,units = "rads",   
 axes = F, border = "black", ticks = T, prop = 3, main = "Sphenisciformes")  
axis.circular(at=circular(sort(seq(0, 11/6\*pi, pi/6), decreasing = T)), c(labels = c("M", "J",  
 "J", "A", "S", "O", "N", "D", "J", "F", "M","A")))  
lines(density(spheni\_circ, bw = 20), col = "red", rotation = "clock", zero = pi/2, shrink = 1.75)



Como pode-se notar pelo gráfico acima, temos uma maior concentração no período de Agosto(A) a Outubro(O). Podemos realizar teste de hipóteses para averiguar se, de fato, há uma diferença na distribuição dos dados ao longo do ano.

## Testes de Hipóteses  
  
rao.spacing.test(spheni\_circ)

##   
## Rao's Spacing Test of Uniformity   
##   
## Test Statistic = 355.7009   
## P-value < 0.001   
##

rayleigh.test(spheni\_circ)

##   
## Rayleigh Test of Uniformity   
## General Unimodal Alternative   
##   
## Test Statistic: 0.832   
## P-value: 0

watson.test(spheni\_circ, dist = "uniform")

##   
## Watson's Test for Circular Uniformity   
##   
## Test Statistic: 712.2958   
## P-value < 0.01   
##

Como podemos notar pelos testes acima, rejeitamos a hipótese de os dados estarem distribuídos uniformemente ao longo do ano.

## Estimativa do Erro Padrão Relativo

#-------------#  
# Erro Padrão #  
#-------------#  
  
# Geral  
n = nrow(data)  
erro\_ab = sd(data$abundance)/sqrt(n)  
erro\_ab\_rel = erro\_ab/mean(data$abundance) \* 100 # em percentual  
erro\_bio = sd(data$biomass)/sqrt(n)  
erro\_bio\_rel = erro\_bio/mean(data$biomass) \* 100 # em percentual  
  
# Mysticeti  
n = nrow(Mysticeti)  
erro\_ab = sd(Mysticeti$abundance)/sqrt(n)  
erro\_ab\_rel = erro\_ab/mean(Mysticeti$abundance) \* 100 # em percentual  
erro\_bio = sd(Mysticeti$biomass)/sqrt(n)  
erro\_bio\_rel = erro\_bio/mean(Mysticeti$biomass) \* 100 # em percentual  
  
# Odontoceti  
n = nrow(Odontoceti)  
erro\_ab = sd(Odontoceti$abundance)/sqrt(n)  
erro\_ab\_rel = erro\_ab/mean(Odontoceti$abundance) \* 100 # em percentual  
erro\_bio = sd(Odontoceti$biomass)/sqrt(n)  
erro\_bio\_rel = erro\_bio/mean(Odontoceti$biomass) \* 100 # em percentual  
  
# Pinipedia  
n = nrow(Pinipedia)  
erro\_ab = sd(Pinipedia$abundance)/sqrt(n)  
erro\_ab\_rel = erro\_ab/mean(Pinipedia$abundance) \* 100 # em percentual  
erro\_bio = sd(Pinipedia$biomass)/sqrt(n)  
erro\_bio\_rel = erro\_bio/mean(Pinipedia$biomass) \* 100 # em percentual  
  
# Procellariiformes  
n = nrow(Procellariiformes)  
erro\_ab = sd(Procellariiformes$abundance)/sqrt(n)  
erro\_ab\_rel = erro\_ab/mean(Procellariiformes$abundance) \* 100 # em percentual  
erro\_bio = sd(Procellariiformes$biomass)/sqrt(n)  
erro\_bio\_rel = erro\_bio/mean(Procellariiformes$biomass) \* 100 # em percentual  
  
# SeaTurtles  
n = nrow(SeaTurtles)  
erro\_ab = sd(SeaTurtles$abundance)/sqrt(n)  
erro\_ab\_rel = erro\_ab/mean(SeaTurtles$abundance) \* 100 # em percentual  
erro\_bio = sd(SeaTurtles$biomass)/sqrt(n)  
erro\_bio\_rel = erro\_bio/mean(SeaTurtles$biomass) \* 100 # em percentual  
  
# Sphenisciformes  
n = nrow(Sphenisciformes)  
erro\_ab = sd(Sphenisciformes$abundance)/sqrt(n)  
erro\_ab\_rel = erro\_ab/mean(Sphenisciformes$abundance) \* 100 # em percentual  
erro\_bio = sd(Sphenisciformes$biomass)/sqrt(n)  
erro\_bio\_rel = erro\_bio/mean(Sphenisciformes$biomass) \* 100 # em percentual  
  
# CRIAR TABELA PARA MOSTRAR  
  
#--------#  
# IC 95% #  
#--------#  
  
# Geral  
n = nrow(data)  
IC95\_ab <- paste0("(",mean(data$abundance)-qt(.975,n-1)\*sd(data$abundance)/sqrt(n)," , ",mean(data$abundance)+qt(.975,n-1)\*sd(data$abundance)/sqrt(n),")")  
IC95\_bio <- paste0("(",mean(data$biomass)-qt(.975,n-1)\*sd(data$biomass)/sqrt(n)," , ",mean(data$biomass)+qt(.975,n-1)\*sd(data$biomass)/sqrt(n),")")  
  
# Mysticeti  
n = nrow(Mysticeti)  
IC95\_ab <- paste0("(",mean(Mysticeti$abundance)-qt(.975,n-1)\*sd(Mysticeti$abundance)/sqrt(n)," , ",mean(Mysticeti$abundance)+qt(.975,n-1)\*sd(Mysticeti$abundance)/sqrt(n),")")  
IC95\_bio <- paste0("(",mean(Mysticeti$biomass)-qt(.975,n-1)\*sd(Mysticeti$biomass)/sqrt(n)," , ",mean(Mysticeti$biomass)+qt(.975,n-1)\*sd(Mysticeti$biomass)/sqrt(n),")")  
  
# Odontoceti  
n = nrow(Odontoceti)  
IC95\_ab <- paste0("(",mean(Odontoceti$abundance)-qt(.975,n-1)\*sd(Odontoceti$abundance)/sqrt(n)," , ",mean(Odontoceti$abundance)+qt(.975,n-1)\*sd(Odontoceti$abundance)/sqrt(n),")")  
IC95\_bio <- paste0("(",mean(Odontoceti$biomass)-qt(.975,n-1)\*sd(Odontoceti$biomass)/sqrt(n)," , ",mean(Odontoceti$biomass)+qt(.975,n-1)\*sd(Odontoceti$biomass)/sqrt(n),")")  
  
# Pinipedia  
n = nrow(Pinipedia)  
IC95\_ab <- paste0("(",mean(Pinipedia$abundance)-qt(.975,n-1)\*sd(Pinipedia$abundance)/sqrt(n)," , ",mean(Pinipedia$abundance)+qt(.975,n-1)\*sd(Pinipedia$abundance)/sqrt(n),")")  
IC95\_bio <- paste0("(",mean(Pinipedia$biomass)-qt(.975,n-1)\*sd(Pinipedia$biomass)/sqrt(n)," , ",mean(Pinipedia$biomass)+qt(.975,n-1)\*sd(Pinipedia$biomass)/sqrt(n),")")  
  
# Procellariiformes  
n = nrow(Procellariiformes)  
IC95\_ab <- paste0("(",mean(Procellariiformes$abundance)-qt(.975,n-1)\*sd(Procellariiformes$abundance)/sqrt(n)," , ",mean(Procellariiformes$abundance)+qt(.975,n-1)\*sd(Procellariiformes$abundance)/sqrt(n),")")  
IC95\_bio <- paste0("(",mean(Procellariiformes$biomass)-qt(.975,n-1)\*sd(Procellariiformes$biomass)/sqrt(n)," , ",mean(Procellariiformes$biomass)+qt(.975,n-1)\*sd(Procellariiformes$biomass)/sqrt(n),")")  
  
# SeaTurtles  
n = nrow(SeaTurtles)  
IC95\_ab <- paste0("(",mean(SeaTurtles$abundance)-qt(.975,n-1)\*sd(SeaTurtles$abundance)/sqrt(n)," , ",mean(SeaTurtles$abundance)+qt(.975,n-1)\*sd(SeaTurtles$abundance)/sqrt(n),")")  
IC95\_bio <- paste0("(",mean(SeaTurtles$biomass)-qt(.975,n-1)\*sd(SeaTurtles$biomass)/sqrt(n)," , ",mean(SeaTurtles$biomass)+qt(.975,n-1)\*sd(SeaTurtles$biomass)/sqrt(n),")")  
  
# Sphenisciformes  
n = nrow(Sphenisciformes)  
IC95\_ab <- paste0("(",mean(Sphenisciformes$abundance)-qt(.975,n-1)\*sd(Sphenisciformes$abundance)/sqrt(n)," , ",mean(Sphenisciformes$abundance)+qt(.975,n-1)\*sd(Sphenisciformes$abundance)/sqrt(n),")")  
IC95\_bio <- paste0("(",mean(Sphenisciformes$biomass)-qt(.975,n-1)\*sd(Sphenisciformes$biomass)/sqrt(n)," , ",mean(Sphenisciformes$biomass)+qt(.975,n-1)\*sd(Sphenisciformes$biomass)/sqrt(n),")")  
  
  
# CRIAR TABELA PARA MOSTRAR