Trabalho Nº6 - Modelos Markovianos

Rodrigo Albanas | Lais Hoffman | Willian Meira

14 de Setembro de 2020

1. Resumo

HI-SEAS (Hawai'i Space Exploration Analog and Simulation) é um habitat em um local isolado semelhante a Marte no lado Mauna Loa da área de sela na Ilha Grande do Havaí, a aproximadamente 8200 pés acima do nível do mar. O HI-SEAS é único, além de sua configuração em um ambiente analógico diferenciado, como:

selecionamos a tripulação para atender às nossas necessidades de pesquisa (em análogos fortuitos, como estações antárticas, os critérios de seleção da tripulação não são controlados pelos pesquisadores); as condições (habitat, missão, comunicações, etc.) são explicitamente projetadas para serem semelhantes às de uma missão de exploração planetária; o local é acessível durante todo o ano e tem muito pouca variação de clima, permitindo estudos de ambiente isolado e confinado de maior duração do que em outros locais; o ambiente semelhante a Marte fornece tarefas analógicas de alta fidelidade, como trabalho de campo geológico realizado por exploradores humanos e / ou robôs. # 2. Introdução ******

1. Base de Dados

1.1 Descrição dos dados

A base original foi retirada do portal *Kaggle* e pode ser obtida no link www.kaggle.com. São dados coletados na estação meteorológica HI-SEAS de setembro a dezembro de 2016, no Hawai

UNIXTime: total de segundos desde 1 de janeiro de 1970; Data: data da coleta no formato "mm/dd/yyyy"; Time: hora da coleta no formato "hh:mm:ss"; WindDirectionG: direção do vento em graus; TimeSunRise e TimeSunSet: hora do nascer e pôr do sol; Radiation: nível de radiação solar medida em W/m2 (watts por metro quadrado); Temperature: temperatura em graus Fahrenheit; Pressure: pressão barométrica medida em mm Hg (milímetros de mercúrio); Humidity: humidade relativa do ar em percentual; Speed: velocidade do vento medida em m/h (milhas por hour);

Para a nossa análise fizemos alguns ajustes na base original. A categorizamos a varíavel referente a direção do ventocriamos algumas covaríaveis com base nos dados originais

TimeHora: DiaNoite: WindDirection: DiaMes: Mes: TimeSeg:

Solar radiation: watts per meter^2 Temperature: degrees Fahrenheit Humidity: percent Barometric pressure: Hg Wind direction: degrees Wind speed: miles per hour Sunrise/sunset: Hawaii time

", sendo dados dos EUA, entre 1997-2002, de acidentes de carro relatados pela polícia nos quais há um evento prejudicial (pessoas ou propriedade) e do qual pelo menos um veículo foi rebocado. Os dados são restritos aos ocupantes do banco da frente, incluem apenas um subconjunto das variáveis registradas e são restritos de outras maneiras também.

Esses conjuntos de dados são dados meteorológicos da

Para cada conjunto de dados, os campos são:

Mode :character

A base original possui uma base de dados com 26.217 observações nas 15 variáveis a seguir.

```
## PACOTES USADOS
library(dplyr)
library(tidyr)
library(stringr)
library(ggplot2)
library(corrplot)
library(gamlss)
library(gamlss.data)
library(gamlss.dist)
## Carregando dados
dados = read.csv("SolarPredictionAjustada.csv",sep = ";", dec = ".", header = TRUE)
str(dados)
'data.frame': 32686 obs. of 17 variables:
$ i..UNIXTime : int 1472724008 1472724310 1472725206 1472725505 1472725809 1472726704 1472727006 14
$ Data
                       "01/09/2016 12:00" "01/09/2016 12:00" "01/09/2016 12:00" "01/09/2016 12:00" ...
                : chr
                : chr "00:00:08" "00:05:10" "00:20:06" "00:25:05" ...
$ Time
 $ WindDirectionG: chr
                      "77,27" "153,44" "142,04" "144,12" ...
                       "06:07:00" "06:07:00" "06:07:00" "06:07:00" ...
$ TimeSunRise : chr
$ TimeSunSet
                      "18:38:00" "18:38:00" "18:38:00" "18:38:00" ...
                : chr
 $ Radiation
                : num 2.58 2.83 2.16 2.21 2.25 2.15 2.07 2.06 2.1 2.07 ...
 $ Temperature : int 51 51 51 51 51 51 52 52 52 ...
 $ Pressure
                : num 30.4 30.4 30.4 30.4 30.4 ...
$ Humidity
                : int 103 103 103 103 103 103 103 103 103 ...
 $ Speed
                : num 11.25 9 7.87 18 11.25 ...
$ TimeHour
                : int 000000011...
$ DiaMes
                : int 1 1 1 1 1 1 1 1 1 1 ...
$ TimeSec
                : int 8 310 1206 1505 1809 2704 3006 3304 3607 3905 ...
$ DayNight
                : int 0000000000...
$ WindDirection : chr
                      "E" "SE" "SE" "SE" ...
 $ Month
                : int 999999999 ...
summary(dados)
 i..UNIXTime
                       Data
                                         Time
                                                        WindDirectionG
                                                        Length: 32686
Min.
       :1.47e+09
                                     Length:32686
                   Length: 32686
 1st Qu.:1.48e+09
                   Class : character
                                     Class : character
                                                        Class : character
Median :1.48e+09
                   Mode :character
                                     Mode :character
                                                        Mode :character
Mean
      :1.48e+09
3rd Qu.:1.48e+09
Max.
      :1.48e+09
                    TimeSunSet
TimeSunRise
                                       Radiation
                                                       Temperature
Length: 32686
                   Length: 32686
                                     Min. : 1.1
                                                      Min.
                                                             :34.0
Class : character
                   Class :character
                                     1st Qu.:
                                                1.2
                                                      1st Qu.:46.0
```

Median :

Mean

Mode :character

2.7

: 207.1

Median:50.0

Mean :51.1

```
3rd Qu.: 354.2 3rd Qu.:55.0
                                     Max. :1601.3 Max. :71.0
   Pressure
                                               TimeHour
                                                              DiaMes
                  Humidity
                                Speed
Min. :30.2 Min. : 8
                            Min. : 0.00 Min. : 0.0 Min. : 1.0
 1st Qu.:30.4 1st Qu.: 56
                            1st Qu.: 3.37
                                           1st Qu.: 6.0
                                                         1st Qu.: 9.0
Median: 30.4 Median: 85
                            Median: 5.62
                                           Median :12.0
                                                          Median:16.0
Mean :30.4
             Mean : 75
                            Mean : 6.24
                                            Mean :11.6
                                                          Mean :15.8
                            3rd Qu.: 7.87
3rd Qu.:30.5
               3rd Qu.: 97
                                            3rd Qu.:18.0
                                                          3rd Qu.:23.0
                                           Max. :23.0
Max.
       :30.6
               Max. :103
                            Max. :40.50
                                                          Max. :31.0
   TimeSec
                   DayNight
                               WindDirection
                                                     Month
Min.
      : 1
               Min.
                      :0.000
                               Length: 32686
                                                 Min.
                                                        : 9.0
 1st Qu.:21617
               1st Qu.:0.000
                                                 1st Qu.:10.0
                               Class :character
Median :43230
               Median :0.000
                                                 Median:11.0
                               Mode :character
                                                 Mean :10.5
Mean
      :43278
               Mean :0.477
3rd Qu.:64849
                3rd Qu.:1.000
                                                 3rd Qu.:11.0
Max.
       :86185
                Max.
                      :1.000
                                                 Max. :12.0
## Ajustando base
dados$WindDirection <- as.factor(dados$WindDirection)</pre>
dados$DayNight <- as.factor(dados$DayNight)</pre>
dados$Month <- as.factor(dados$Month)</pre>
dados$Radiation <- as.numeric(dados$Radiation)</pre>
dados$Temperature <- as.numeric(dados$Temperature)</pre>
dados$Pressure <- as.numeric(dados$Pressure)</pre>
dados$Humidity <- as.numeric(dados$Humidity)</pre>
dados$Speed <- as.numeric(dados$Speed)</pre>
dados = dados[,-c(1,2,3,4,5,6)]
str(dados)
'data.frame': 32686 obs. of 11 variables:
$ Radiation
              : num 2.58 2.83 2.16 2.21 2.25 2.15 2.07 2.06 2.1 2.07 ...
$ Temperature : num 51 51 51 51 51 51 52 52 52 ...
 $ Pressure
              : num 30.4 30.4 30.4 30.4 30.4 ...
$ Humidity
              : num 103 103 103 103 103 103 103 103 103 ...
 $ Speed
              : num 11.25 9 7.87 18 11.25 ...
$ TimeHour
             : int 000000011...
$ DiaMes
               : int 111111111...
 $ TimeSec
             : int 8 310 1206 1505 1809 2704 3006 3304 3607 3905 ...
 $ DayNight
               : Factor w/ 2 levels "0", "1": 1 1 1 1 1 1 1 1 1 1 ...
 $ WindDirection: Factor w/ 8 levels "E","N","NE","NW",..: 1 6 6 6 3 1 6 1 1 1 ...
               : Factor w/ 4 levels "9","10","11",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ Month
summary(dados)
```

Radiation Temperature Pressure Humidity Speed

```
1.1
                 Min.
                        :34.0
                               Min.
                                      :30.2
                                              Min. : 8
                                                            Min. : 0.00
1st Qu.:
           1.2
                 1st Qu.:46.0
                               1st Qu.:30.4
                                              1st Qu.: 56
                                                            1st Qu.: 3.37
                Median:50.0
                               Median:30.4
                                                            Median: 5.62
Median :
           2.7
                                              Median: 85
Mean : 207.1
                       :51.1
                               Mean
                                      :30.4
                                                    : 75
                                                            Mean
                                                                  : 6.24
                Mean
                                              Mean
3rd Qu.: 354.2
                 3rd Qu.:55.0
                                3rd Qu.:30.5
                                              3rd Qu.: 97
                                                            3rd Qu.: 7.87
Max.
      :1601.3
                Max.
                       :71.0
                               Max.
                                       :30.6
                                              Max.
                                                     :103
                                                            Max.
                                                                   :40.50
   TimeHour
                   DiaMes
                                 TimeSec
                                             DayNight WindDirection
Min. : 0.0
              Min.
                     : 1.0
                             Min.
                                         1
                                             0:17078
                                                       S
                                                               :9232
1st Qu.: 6.0
              1st Qu.: 9.0
                                             1:15608
                                                       SE
                                                               :7516
                              1st Qu.:21617
Median:12.0
              Median:16.0
                             Median :43230
                                                               :4590
     :11.6
                                    :43278
                                                       Ε
                                                               :4365
Mean
              Mean
                     :15.8
                             Mean
3rd Qu.:18.0
               3rd Qu.:23.0
                              3rd Qu.:64849
                                                       N
                                                               :3076
Max.
      :23.0
              Max.
                     :31.0
                             Max.
                                    :86185
                                                       NW
                                                               :1735
                                                       (Other):2172
Month
9:7417
10:8821
11:8284
12:8164
```

2 Análise Descritiva

2.1 Medidas de Resumo

```
library(gridExtra)

g1 <- ggplot(dados, aes(x=Radiation)) + geom_histogram()+ xlab('Radiation')+ ylab('')

g2 <- ggplot(dados, aes(x=Temperature)) + geom_histogram()+ xlab('Temperature')+ ylab('')

g3 <- ggplot(dados, aes(x=Pressure)) + geom_histogram()+ xlab('Pressure')+ ylab('')

g4 <- ggplot(dados, aes(x=Humidity)) + geom_histogram()+ xlab('Humidity')+ ylab('')

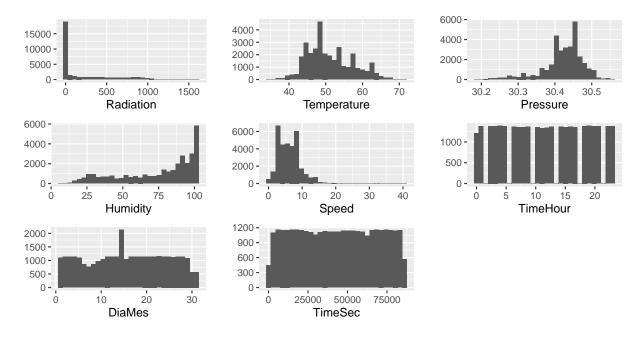
g5 <- ggplot(dados, aes(x=Speed)) + geom_histogram()+ xlab('Speed')+ ylab('')

g6 <- ggplot(dados, aes(x=TimeHour)) + geom_histogram()+ xlab('TimeHour')+ ylab('')

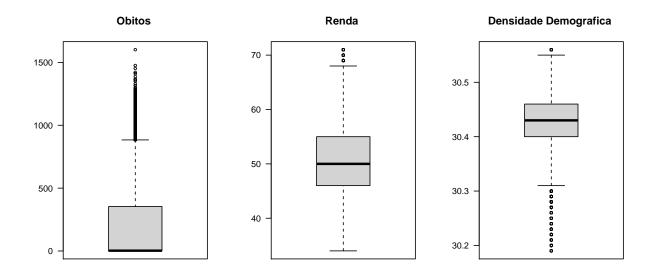
g7 <- ggplot(dados, aes(x=DiaMes)) + geom_histogram()+ xlab('DiaMes')+ ylab('')

g8 <- ggplot(dados, aes(x=TimeSec)) + geom_histogram()+ xlab('TimeSec')+ ylab('')

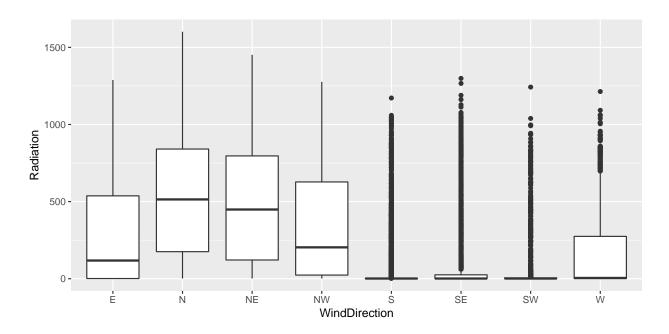
grid.arrange(g1,g2,g3,g4,g5,g6,g7,g8,ncol=3, nrow=3)</pre>
```



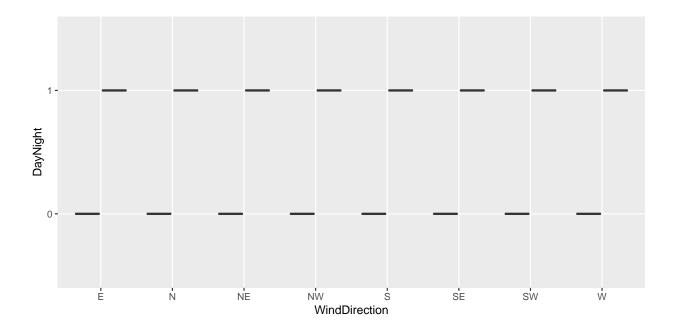
```
x11()
par(mfrow=c(1,3))
boxplot(dados$Radiation, xlab = '', ylab = '', main = 'Obitos ', las=1)
boxplot(dados$Temperature, xlab = '', ylab = '', main = 'Renda ', las=1)
boxplot(dados$Pressure, xlab = '', ylab = '', main = 'Densidade Demografica ', las=1)
```

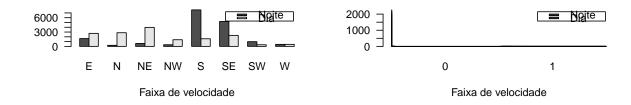


Direção vento X Radiaton

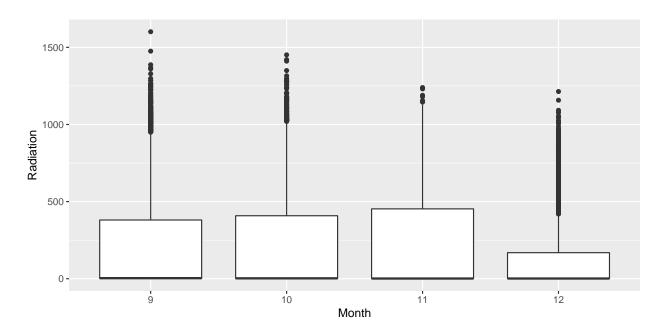


Direção vento X DiaNoite





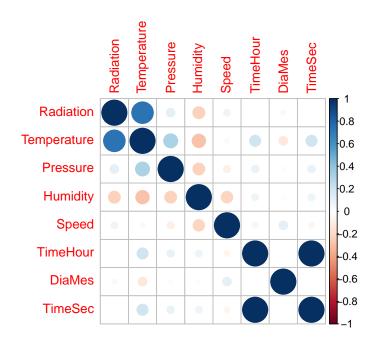
Direção vento X Radiaton



```
X <- cor(dados[,1:8])
K <- cor(dados[,c("Radiation","Temperature")])
X</pre>
```

```
Radiation Temperature Pressure Humidity
                                                    Speed
                                                           TimeHour
            1.000000
                        0.73495 0.11902 -0.22617 0.07363
Radiation
                                                           0.004398
Temperature 0.734955
                        1.00000 0.31117 -0.28505 -0.03146 0.197464
Pressure
            0.119016
                        0.31117 1.00000 -0.22397 -0.08364 0.091069
Humidity
           -0.226171
                       -0.28505 -0.22397 1.00000 -0.21162 0.077899
Speed
            0.073627
                       -0.03146 -0.08364 -0.21162 1.00000 -0.057939
TimeHour
            0.004398
                        DiaMes
            0.039978
                       -0.12371 -0.02463 0.01464 0.11734 -0.008010
TimeSec
            0.004348
                        0.19723 \quad 0.09107 \quad 0.07785 \quad -0.05791 \quad 0.999134
              DiaMes
                      TimeSec
Radiation
            0.039978 0.004348
Temperature -0.123705 0.197227
Pressure
           -0.024633 0.091066
Humidity
            0.014637 0.077851
Speed
            0.117337 -0.057908
TimeHour
           -0.008010 0.999134
            1.000000 -0.007966
DiaMes
TimeSec
           -0.007966 1.000000
```

```
corrplot(X, method="circle")
```



```
X <- cor(dados[,1:8])
K <- cor(dados[,c("Radiation","Temperature")])
X</pre>
```

```
Radiation Temperature Pressure Humidity
                                                   Speed TimeHour
Radiation
            1.000000
                        0.73495 0.11902 -0.22617 0.07363
                                                         0.004398
                        1.00000 0.31117 -0.28505 -0.03146 0.197464
Temperature 0.734955
Pressure
            0.119016
                        0.31117 1.00000 -0.22397 -0.08364
                                                         0.091069
Humidity
           -0.226171
                       -0.28505 -0.22397 1.00000 -0.21162 0.077899
Speed
            0.073627
                       -0.03146 -0.08364 -0.21162 1.00000 -0.057939
TimeHour
            0.004398
                        0.19746 0.09107 0.07790 -0.05794 1.000000
DiaMes
            0.039978
                       -0.12371 -0.02463 0.01464 0.11734 -0.008010
TimeSec
            0.004348
                        DiaMes
                      TimeSec
Radiation
            0.039978 0.004348
Temperature -0.123705 0.197227
Pressure
           -0.024633 0.091066
Humidity
            0.014637 0.077851
Speed
            0.117337 -0.057908
TimeHour
           -0.008010 0.999134
DiaMes
            1.000000 -0.007966
TimeSec
           -0.007966 1.000000
```

```
corrplot(X, method="circle")
```

