Matriculation number:	2	4	5	4	7

Examination Assignment

Module: Data Analysis and Statistics

Exam part: Data Analysis and Statistics

Examiner: Prof. Dr. Schwind, Dipl.-Biol. Ralf Darius

Deadline for the submission: 31.08.2018, 11:59 pm

Study program	Begin of studies	Last name, First name
Information Engineering and Computer Science (M.Sc.)	SS2018	Ali, Md Monsur

Assessment criteria and number of points that can be achieved:

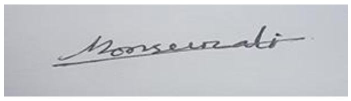
Maximum number of points	Skills and Expertise	Systematic and scientific Quality	Quality of the results	Presentation of the results
100	45	15	30	10

Result:

Points	**Mark**	Skills and Expertise	Systematic and scientific Quality	Quality of the results	Presentation of the results

Statement of Authorship

This report is the result of my own work. Material from the published and unpublished work of others, which is referred to in the report, is credited to the author in the text.



____

Md Monsur Ali

Abstract

This paper is all about killer whale or orca and how natural selection and the shape of other evolutionary forces by the environment of the environment, which pervades their survival and reproduction in the face of environmental challenges posed. Here one thousand female whales are created and analysis them to find out information in different environments. Using histogram in R, it shows the statistics of ages at end of reproduction and number of offspring recruited.

Table of Contents

List of Tables & Figures

1.Introduction:

The main point of life history theory, a branch of developmental ecology, is to clarify the exceptional assorted variety in life accounts among species. The investigation of life history advancement is in this way about understanding adjustment, the most major issue in a developmental environment [1]. It is about the varieties that living beings show regarding the grouping of occasions in their lives identified with survival and proliferation that happen from birth through death. So, it is split among process like growth, reproduction and health. From many years, scientists are observing these processes to find out how ecology life is going on and protect them many dangerous aspects. The killing whales or orca is one of the species, scientists are observing them about whole life duration birth to death. Orcas are frequently called killing whales. Despite the fact that they don't commonly assault people, this name is still well-picked because of the creature's capacity to bring down vast marine creatures, for example, ocean lions and whales. Indeed, orcas will go after any creature they find in the ocean, noticeable all around finished the water or along the coastline [2]. There are up to five particular killer whales. They are separate by their physical appearance, geographical race and subspecies. The female whales become reproductively active at the age of 15 and until they are 40. Calves nurture for around 2 years after birth, yet remain nearly connected with their mom for around four years. Moms keep up a level of parental venture for that period, prompting a between birth interim of five years. Amid those years, calves experience the ill effects of a yearly mortality of 20%. After their fifth year, female calves make due to enlistment with a likelihood of 0.98. Feminine life spans average fifty years however could survive well into their 70-80s in exceptional cases. Males become sexually mature at the age of fifteen however do not generally reproduce till age twenty-one. Male killer whales usually do not live as long as females. within the wild, males average thirty years, with a most of 50–60 years in exceptional cases.

1.1Problem Statement

Investigate how the availability and variability of resources affect the individual fitness, the total reproductive output and the age of female Orcas at death. The investigation seeks to support a better understanding of the effect of the environment on both the fitness of an individual Orca and the reproductive ability of an entire Orca population. The investigation will be done by means of a simulation programmed in R.

1.2Tasks Covered

The numbered list below to step by step build and run the simulation in R. The list consists of different parts were done one after another. The first three sections will cover the modeling part:

i. Modelling a single female without calf.
ii. Modelling a single female with calf.
iii. Modelling several females.
iv. Experiencing and simulating different environmental scenarios. The scenarios are covered in the following sections:

    a) Scenario 1: Rich and stable environment
    b) Scenario 2: Poor and stable environment
    c) Scenario 3: Rich and variable environment
    d) Scenario 4: Poor and variable environment

2.Methods and Materials:

The tasks are completed using several different methods in RStudio. Some of the methods are explained below:

2.1RStudio:

R is a free and open source software program. R is used for statistical analysis based on script language. The main advantages of R are the fact that R is freeware and that there is a lot of help available online[3]. It is quite similar to other programming packages such as MatLab. RStudio is open source IDE for R. User can clearly view graphs, data tables, R code, and output all at the same time by using rstudio. User can import various of file on this platform and analysis like CSV, Excel, SAS (*.sas), SPSS (*.sav), and Stata (*.dta) files. RStudio have lots of open source libraries, ones can download and install the libraries and work. It is available for windows, Mac and Linux user[3].

2.2Generating random numbers:

R has many functions to generate a random number. Such as uniform distribution, binomial distribution, normal distribution etc. The full rundown of standard distribution accessible can be seen utilizing? distribution. Function that generate random numbers begin with the letter r. Like rnorm() generates from a normal distribution, runif() generates from the continuous uniform distribution. Here a short description is given of generating random numbers:

i. **Normal distribution:** A random number can be generated from a normal distribution using rnorm() function[4]. Here we specify a sample number to be generated. We also specify mean and standard deviation of the distribution. The distribution defaults to 0 mean and 1 standard deviation, if not given.

    > rnorm(3)   # generates 3 random number

[1] -1.1383656   0.2016713   -0.4602043

> rnorm(3, mean=10, sd=2)   # provide our own mean and standard deviation

[1] 9.856933   9.024286   10.822507

ii. **Uniform distribution:** A random number can be generated from a continuous distribution using runif() function[5]. We need to specify how many numbers we want to generate. Additionally, we can specify the range of the uniform distribution using max and min argument. If not provided, the default range is between 0 and 1.

> runif(3) # generates 3 random number

[1] 0.8090284   0.1797232   0.6803607

> runif(3, min=5, max=10)   # define the range between 5 and 10

[1] 7.099781   8.355461   5.173133

iii. **Binomial distribution:** A random number can be generated from a binomial distribution using rbinom() function[6]. We need to specify how many numbers we want to generate. Additionally, there are then two parameters n (the number of Bernoulli trials) and p (the success probability).

> rbinom(1,n,p)   # 6 successes in 10 trials (n=10, p=0.5)

[1] 6

> rbinom(5,n,p)   # 5 binomial number

[1] 6 6 4 5 4

iv. **Poisson distribution:** A random number can be generated from a poision distribution using rpois() function[7]. Additionally, there are then one parameter a poisson distribution rate.

>rpois(10, 5.3)   # 10 poisson numbers with 5.3 rate

[1] 6 4 9 4 5 5 5 4 3 5

2.3Conditional statements:

Conditionals are articulations that perform diverse calculations or activities relying upon whether a predefined Boolean condition is TRUE or FALSE[8]. Basic mathematic conditions are shown below:

> 5<3   # Check the condition

[1] False

> 5==5   # Check the condition

[1] True

> 2+2=4 && 3-1=1   #Check both condition

[1] False

>2*3=5 || 6/3=2   # Check both condition

[1] True

So, here first one showed that 5 is less than 3 which is False and second one is checking 5 is equal to 5 which is True. Last two example showed two condition of each example. And operator is true when both conditions are true. Besides OR operator is true when one condition is true.

- **The if() Statement:**

If() statement is common in every programming language. It performs like the given condition is true then the entire block codes are executed otherwise stop[8]. Here is structure of if() statement:

```
If(condition)
{
commands
}
```

The following if statement first examine the condition. The condition is true then it execute the entire block code.

```
> If(5>4)
{
Print("5 is bigger than 4")
}
[1] 5 is bigger than 4
```

- **The if() and else if() Statement:**

In this statement, first examine the first condition if first condition is true then execute its block commands or it examine other if else() statement[8]. All statement condition is false then it execute only else command otherwise stop.

```
>k = 100
if(k > 100){
print("Greater than 100")
} else if (k < 100){
print("Less than 100")
} else {
print ("Equal to 100")
}
[1] Equal to 100
```

2.4Iteration and loop:

Iteration is the way to repeat the condition. Using iteration, R supports vectorization. Iterative tasks require a different programming device, called a loop, which comprises of two parts[9].

1. a loop declaration and
2. the main body of the loop.

Here, we will introduce two types of loops, the for loop and while loop.

- **Loop For:**

A for loop repeats until a specified condition evaluates to false. In R, for loop looks like this:

```
for(counter in min:max)   # This line declares the loop counter and range.
  {
  The main body of loop - the commands to be repeated.
  }
```

The following for loop explains, it repeats until the condition of loop is over. A variable year store 2010 to 2015 years and the main body is printed the commands until the years are over.

```
>for (year in 2010:2015)
{
print(paste("The year is", year))
}
[1] "The year is 2010"
[1] "The year is 2011"
[1] "The year is 2012"
[1] "The year is 2013"
[1] "The year is 2014"
[1] "The year is 2015"
```

- **While Loop :**

A while loop executes its statements as long as a specified condition evaluates to true. In R, while loop looks like this:

```
while (conditions) # This line declares the loop conditions.

  {

 # The main body of loop - the commands to be repeated.

   }
```

The following code describe the while loop. The loop is repeated until count is less than 5. On main body it prints "Hello", "while loop"

```
> v <- c("Hello", "while loop")
count <- 2
while (count < 4)
{
print(v)
count = count + 1
}
[1] "Hello" "while loop"
[1] "Hello" "while loop"
```

3.Analysis and Result

3.1 Modelling

First of all, before getting to the scenarios and generating respective histograms, we need to model the calves using different attributes. Modelling is an object's representation so that it can be used for simulation purposes later to predict its future state. In this case, we have models three different objects:

i. Modelling a single female without a calf.
ii. Modelling a single female with calf.
iii. Modelling several females.

Their respective R code can be viewed at the end of this report.

3.2 Simulation

After modelling is complete, simulation techniques are applied to predict an object's future. Simulation is functioning of a system to analyse the interaction of individual elements with each other and their effect on the environment. In this case four different environments are considered which are explained below.

i. Scenario 1: Simulating rich and stable environment.
ii. Scenario 2: Simulating poor and stable environment.
iii. Scenario 3: Simulating rich and variable environment.
iv. Scenario 4: Simulating poor and variable environment.

In this case, the modelled calves are put under these environments and their interaction with each other along with their effect on each environment is noted. The respective results are mentioned in the next section.

3.3Results:

In this section, all the results after simulating the modelled calves in four different environments are derived and discussed:

3.3.1Rich and stable environment

We will interpret the number of recruits produced by each female as its inclusive fitness. In rich and stable environment parameters are set as, enMu= 20 and enSD= 1. To visualize this date a simple hist() command is used, which shows a histogram between the two values. The following two figures are displayed.

The value of average recruits is 0.1445 in the rich and stable environment. The below histogram in figure 1.1 depicts that the distribution of ages at the end of reproduction of 1000 whales in rich and stable environment. More than 800 whales are losing their reproduction capacity between the ages 40 to 43 years. Almost 130 whales are also losing reproductivity between the ages14 to 16 years. After age of 16 years, a few numbers of whales lose reproduction ability between 16 to 20 years of age.

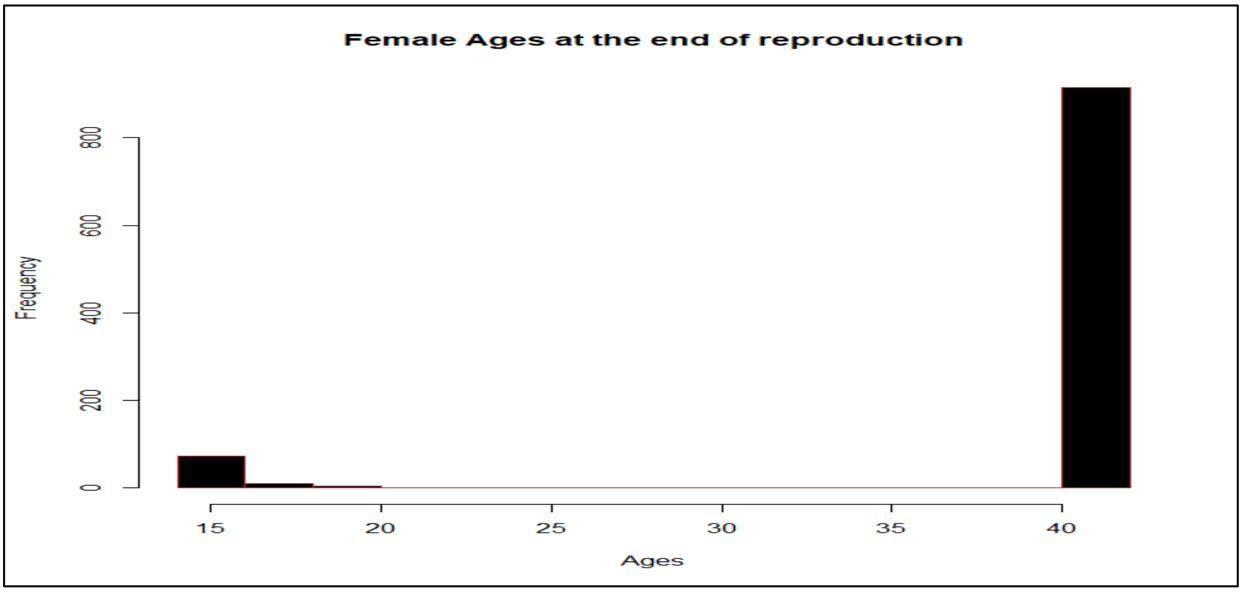


Figure 1.1: Ages in rich and stable environment

The figure 1.2 below shows the histogram of inclusive fitness (number of offspring's recruited) among 1000 whales in the rich and stable environment. It can be seen from the figure that more than 700 whales do not recruit any offspring's whereas more than 200 females recruit a maximum of one offspring between 15 to 40 years of age. A calf is considered an offspring when it is 15 years of age. However, in rich and stable environment, majority of female whales fail to produce any offspring.

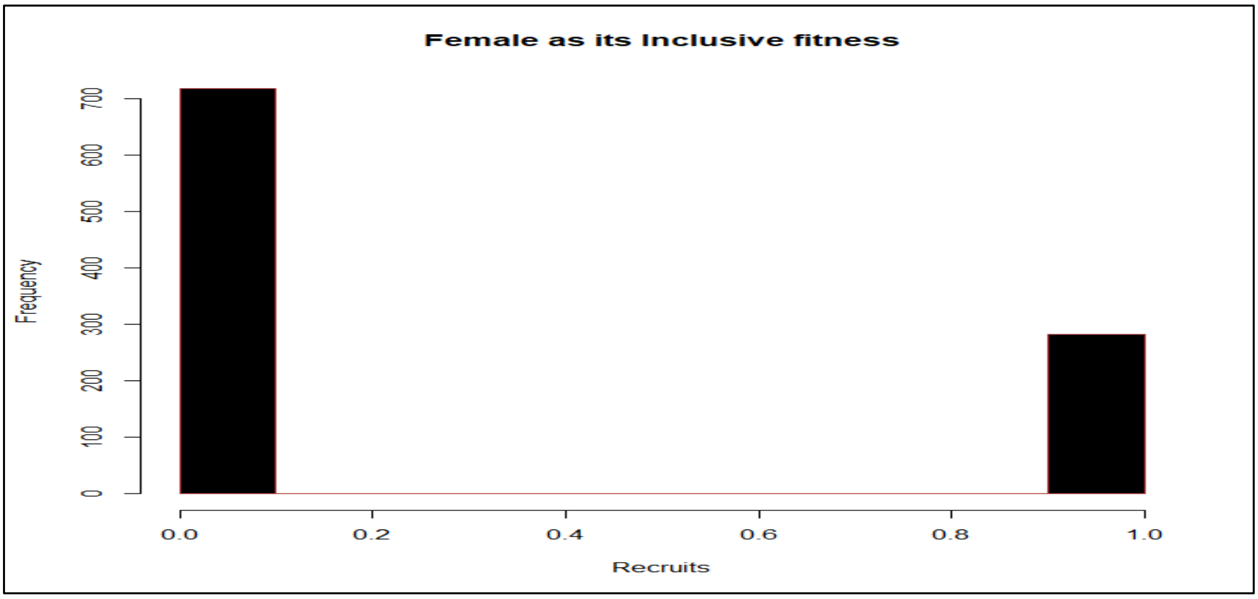


Figure 1.2: Recruits in rich and stable environment

3.3.2Poor and stable environment:

We will interpret the number of recruits produced by each female as its inclusive fitness. In rich and stable environment parameters are set as, enMu= 2 and enSD= 1. Using hist() command the following figure are showed:

The value of average recruits is 0.002 in the poor and stable environment. The below histogram in figure 2.1 depicts that the distribution of ages at the end of reproduction of 1000 whales in poor and stable environment. At beginning of the ages, more than 100 whales are losing reproduction capacity. After the age of 22 years, the number of whale reproduction is decreasing dramatically 100 to 10 whales. It increases 20 whales at the ages between 40 to 43 years.

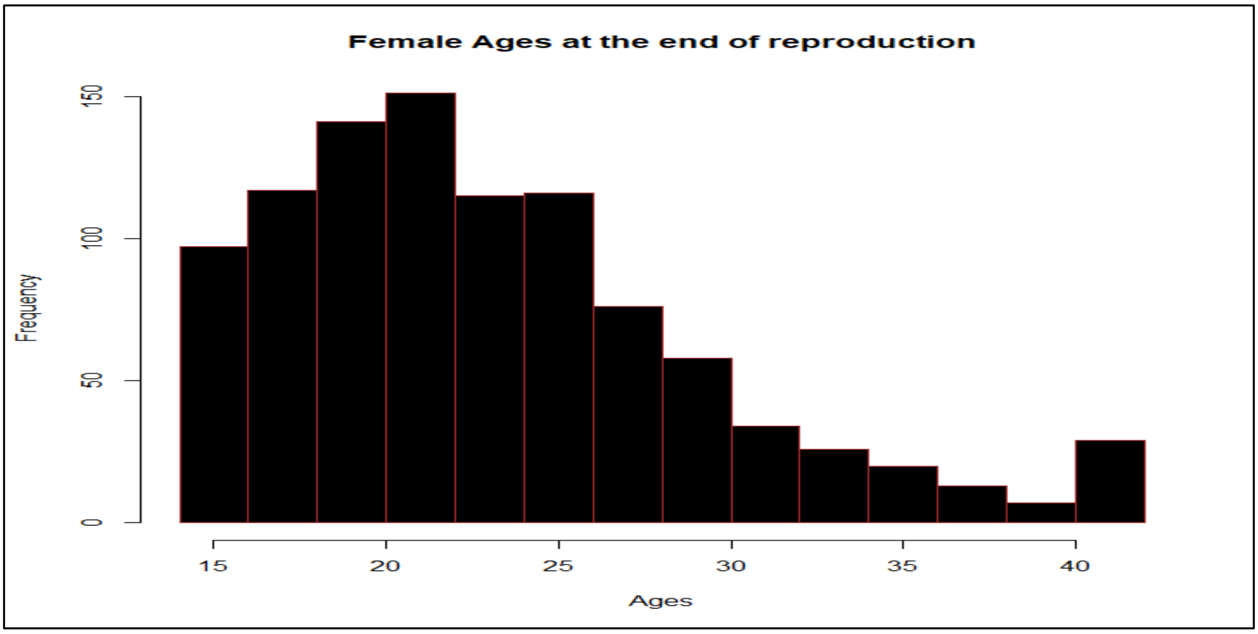


Figure 2.1: Ages in poor and stable environment

The figure 2.2 below shows the histogram of inclusive fitness (number of offspring's recruited) among 1000 whales in the poor and stable environment. It can be seen from the figure that more than 900 whales do not recruit any offspring's whereas just a few numbers of females recruit a maximum of one offspring between 15 to 40 years of age. A calf is considered an offspring when it is 15 years of age. However, in poor and stable environment, almost all the female whales fail to produce any offspring

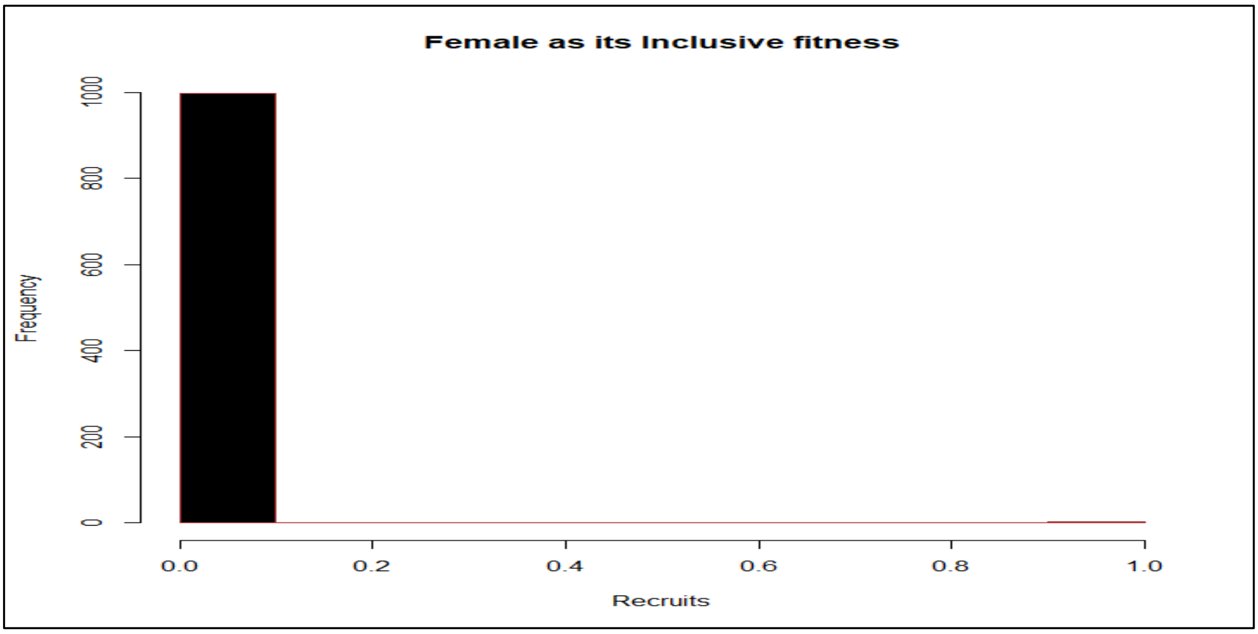


Figure 2.2: Recruits in poor and stable environment

3.3.3Rich and variable environment:

We will interpret the number of recruits produced by each female as its inclusive fitness. In rich and stable environment parameters are set as, enMu= 20 and enSD= 30. Using hist() command the following figure are showed:

The value of average recruits is 0.224 in the poor and stable environment. The below histogram in figure 3.1 depicts that the distribution of ages at the end of reproduction of 1000 whales in rich and variable environment. It is showed that more than 700 female whale reproduction is

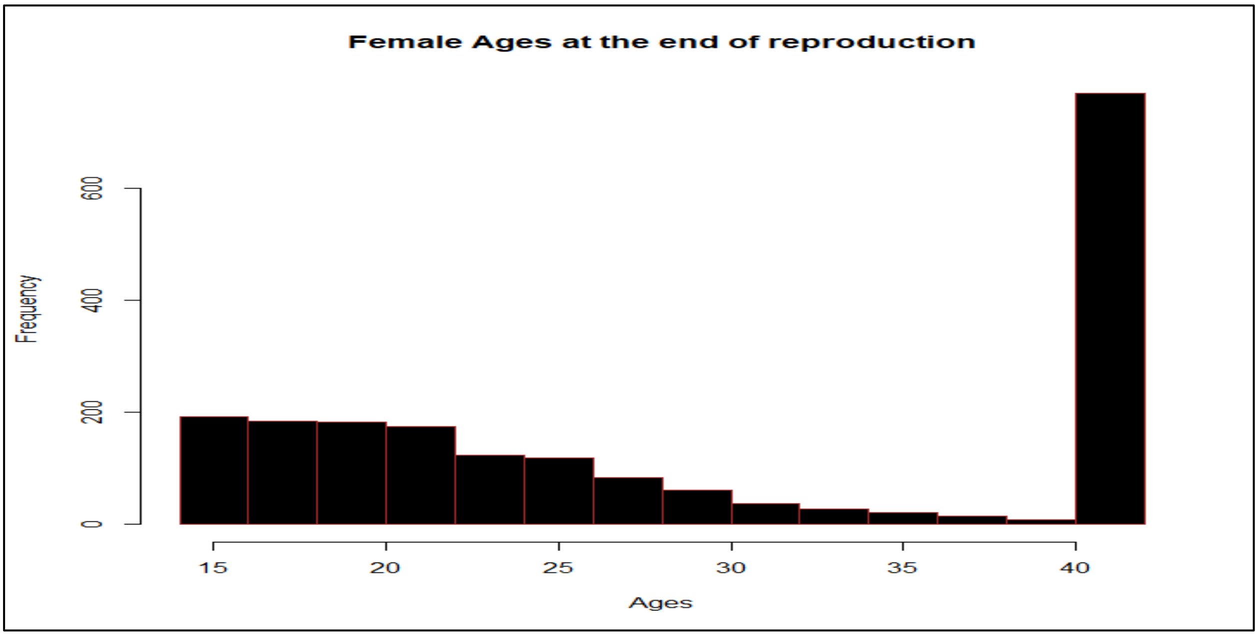


Figure 3.1: Ages in rich and variable environment

end at the ages between 40 to 43 years. From the age of 14 years, the number of whale reproduction are decreasing frequently 195 to 5 until the age of 40 years.

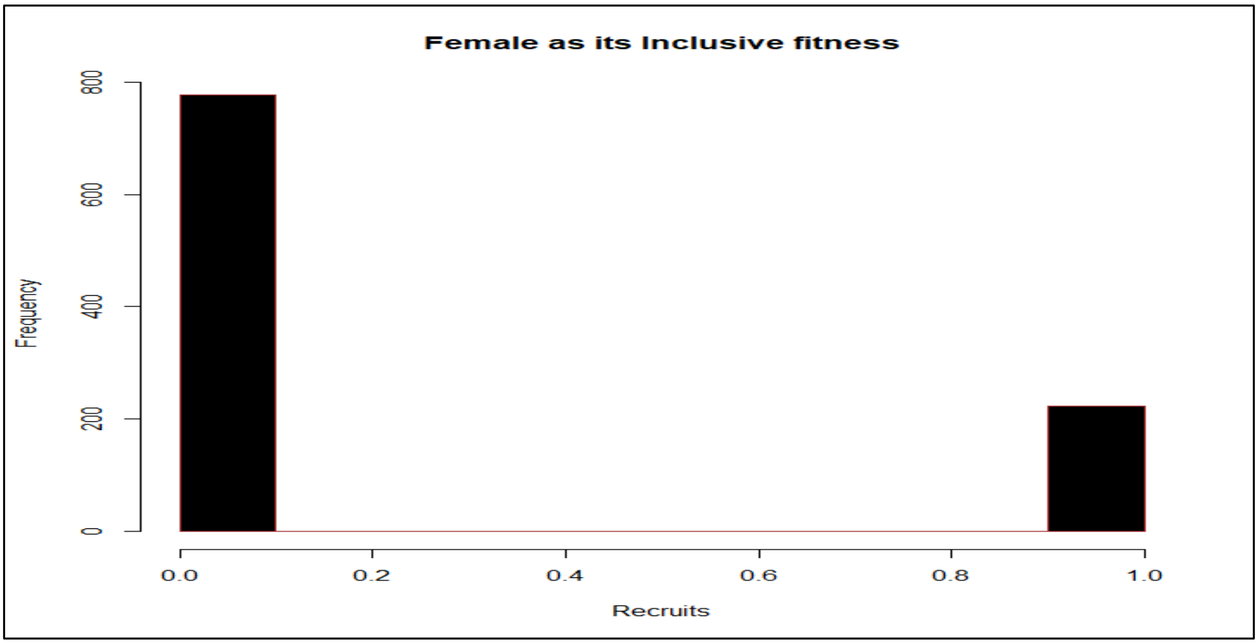


Figure 3.4: Recruits in rich and variable environment

The figure 3.4 upper shows the histogram of inclusive fitness (number of offspring's recruited) among 1000 whales in the rich and variable environment. It can be seen from the figure that about 800 whales do not recruit any offspring's whereas just 200 females recruit a maximum of one offspring between 15 to 40 years of age. A calf is considered an offspring when it is 15 years of age. However, in rich and variable environment just like the rich and stable environment majority of the female whales fail to produce any offspring's.

3.3.4Poor and variable environment:

We will interpret the number of recruits produced by each female as its inclusive fitness. In rich and stable environment parameters are set as, enMu= 2 and enSD= 30. Using hist() command the following figure are showed:

The value of average recruits is 0.028 in the poor and stable environment. The below histogram in figure 4.1 depicts that the distribution of ages at the end of reproduction of 1000 whales in rich and variable environment. The ages between 14 to 16 years, the number of whale reproduction is 145. Then one years later, the number increase to 250. A frequently decreasing is happening from the ages between 19 to 40 years and the number of whales with the end of reproduction is between 145 to 20. At the age of 45 years, it increases to 130.

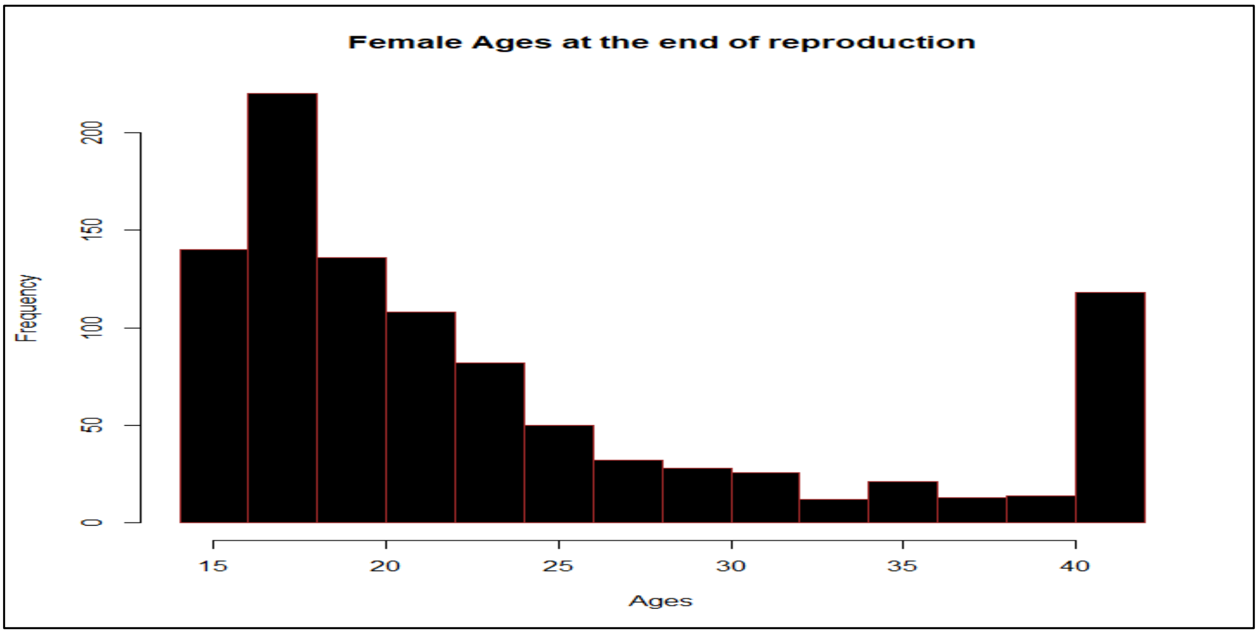


Figure 4.1: Ages in poor and variable environment

The figure 4.2 below shows the histogram of inclusive fitness (number of offspring's recruited) among 1000 whales in the poor and variable environment. It can be seen from the figure that more than 900 whales do not recruit any offspring's whereas just a few numbers of females recruit a maximum of one offspring between 15 to 40 years of age. A calf is considered an offspring when it is 15 years of age. However, in poor and variable environment, almost all the female whales fail to produce any offspring.

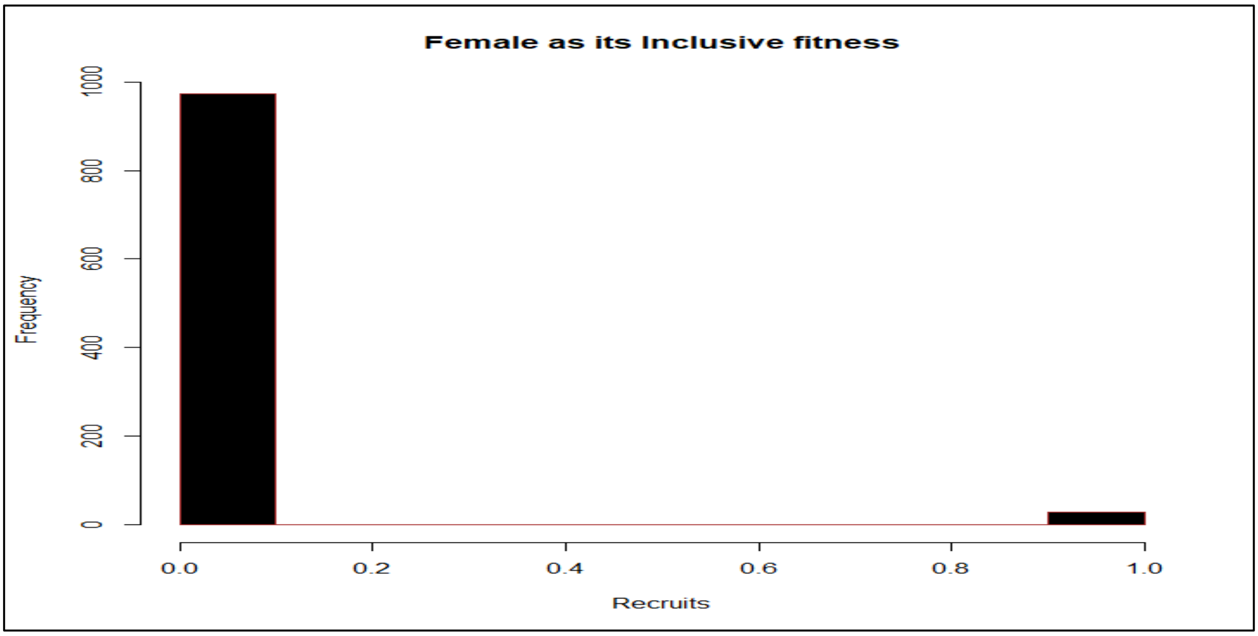


Figure 4.2: Recruits in poor and variable environment

4.Discussion:

This report observing the killer whale lifespan, productivity, their calf birth and their survival. Some analyses were done on how age and environment are effective on the life of female whale breeding. After modelling and simulation, careful analysis was made and we deduced the following results:

When the environment is rich and stable or rich and variable, the maximum of female whale ends their productivity after the age of 40 years. Same as inclusive fitness for both environments is rich and stable or rich and variable. The maximum number of whales do not recruit any offspring's whereas just 200 females recruit a maximum of one offspring between 15 to 40 years of age.

When the environment is poor and stable or poor and variable, the maximum number of whales between 14 and 16 years old lose their reproductive abilities when exposed to poor and stable environment. In stable environment maximum number of whales between ages 20 to 25 lose their reproductive capacity. On the other hand, in the poor and variable environment maximum number of whales lose productivity between 15 to 25 years however there is a linear decrease decreased linearly to till 38 years and a sharp rise after the age of 40 years.

The maximum average number of recruits produced by each female is 0.224 for rich and variable environment. Besides, the lowest number of recruits produced by each female is 0.002 for poor and stable environment.

As should be obvious from the histograms, when these killer whales are analysed under various environment conditions, their reactions change every once in a while.

Reference

1. https://www.nature.com/scitable/knowledge/library/life-history-evolution-68245673
2. https://www.livescience.com/27431-orcas-killer-whales.html
3. https://cran.r-project.org/doc/contrib/Torfs+Brauer-Short-R-Intro.pdf
4. http://mathworld.wolfram.com/UniformDistribution.html
5. https://en.wikipedia.org/wiki/Normal_distributi
6. https://en.wikipedia.org/wiki/Probability_distribut
7. https://en.wikipedia.org/wiki/Binomial_distribution
8. https://en.wikipedia.org/wiki/Conditional_(computer_programming)
9. https://en.wikipedia.org/wiki/Iteration
10. https://developer.mozilla.org/enUS/docs/Web/JavaScript/Guide/Loops_and_iteration

Appendix

RStudio code are shown below:

```r
library(sqldf)
library(dplyr)
library(tidyverse)
library(dplyr)
library(ggplot2)

#Md Monsur Ali
#Matriculation No: 24547
#Data Analysis and Statistics

#Modelling a single female without a calf and Modelling a single female with a calf: :

#Ques:16
recruits <- c()
ages<- c()

#Ques:17
#Simulate 1000 female life histories
for (i in 1:1000){

#Define all variables
#Ques:1
co<-rnorm(1,100,10)

#Ques:2
age<- 15

#Ques:3
alive <- 1

#Ques:4
enMu<- 0
enSD<- 20

```

Figure 5.1: Commented RStudio code 1

```r
#------------Ques 18-24  ------------#

# #Scenario 1: Rich and stable environment
# enMu<- 20
# enSD<- 1

# #Scenario 2: Poor and stable environment
# enMu<- 2
# enSD<- 1
#
# #Scenario 3: Rich and variable environment
# enMu<- 20
# enSD<- 30
#
# #Scenario 4: Poor and variable environment
# enMu<- 2
# enSD<- 30

#------------Ques 18-24  ------------#

#Ques:5
sr<-0.8

#Ques:6
s0<- -2
s1<- 0.05

#Ques:10
calf <- 0
calfage <- 0
offspring <- 0

#Ques:12
b0 <- -10
b1 <- 0.1

```

Figure 5.2: Commented RStudio code 2

```r
#Ques:13
inv <- 10
recruit <- 0
calfborn <- c()


#Ques:7
while (alive>0 && age<=40 ) {

  #Ques:9
  sr <- exp(s0 + s1 * co)/(1 + exp(s0 + s1 * co))
  if (rbinom(1, 1, sr) == 1) {

    #Modelling a single female with a calf:
    #female doesn't have a calf,
    #Ques:14
    if (calf == 0) {
      b <- exp(b0 + b1 * co)/(1 + exp(b0 + b1 * co))
      if (rbinom(1, 1, b) == 1 ) {

        #calf is born
        calf <- 1

      }
    } else {

      #female has a calf
      co <- co - inv
      if (rbinom(1, 1, 0.8) == 1) {
        #calf is going to surviv
        calfage <- calfage + 1

      } else {

        #calf has died
        calf <- 0
        calfage <- 0
      }
```

Figure 5.3: Commented RStudio code 3

```r
        calf <- 0
        calfage <- 0
      }
      #15 independence reached
      if (calfage > 4) {
        offspring <- offspring + 1
        calfborn <- cbind(calfborn,age)
        calf <- 0
        calfage <- 0
      }
    }
    age <- age + 1
    co <- co + rnorm(1, enMu, enSD)
  } else {

    # female is dead
    alive <- 0
  }
}

if (length(calfborn)>0){
  for (j in 1:length(calfborn)){
    if ((age-calfborn[[j]]-5)>=15 && rbinom(1,1,0.98)){
      recruit <- recruit + 1
    }
  }

}


ages <- cbind(ages,age)
recruits <- cbind(recruits,recruit)
avgrecruits <- mean(recruits)
}
```

Figure 5.4: Commented RStudio code 4

```r
avgrecruits <- mean(recruits)
}

# Visualizing results

#Ques:19
hist(ages, main = "Female Ages at the end of reproduction", xlab = "Ages", border="brown", col="black")

#Ques:20
hist(recruits, main = "Female as its Inclusive fitness", xlab = "Recruits", border="brown", col="black")

#Ques:21
avgrecruits
```

Figure 5.5: Commented RStudio code 5