

Tarea diplomado: Analisis de caso: Mortalidad en MarineFjordland en 2021

Carolina San Martin

2022-06-02

Titulo: “Analisis de caso, mortalidad en MarineFjordland en 2021”

Descripción del problema a resolver

En 2021, en Marinefjordland, un zona de fiordos con 5 areas de manejo (simil a un ‘barrio’ en Chile) donde se cultiva salmón del atlantico, se vieron afectados 87 de 150 centros de cultivo. El problema fue una mortalidad significativa de inicio repentino, comenzando en un nivel bajo, pero aumentando rápidamente hasta el 5% de la población en riesgo por día en los sitios afectados. Los datos fueron recopilados en 2021, sin resultados concluyentes, no fue posible determinar el patogeno que lo causó, ni aislarlo. Por lo tanto la definición de caso fue: mortalidad significativa de inicio súbito que comienza en un nivel bajo, pero aumenta rápidamente al 5% o más de la población en riesgo por día en los sitios afectados. El centro de cultivo se considera caso (1) cuando las mortalidades alcanzan el 5% por día y no caso (0). Se probaron varios enfoques terapéuticos, pero ninguno ayudo significativamente a evitar la progresión de la enfermedad (aumento de mortalidad). Se cosecharon algunos de los sitios afectados, en la medida que la talla se los permitió, 5 sitios sacrificaron el stock, y en el resto, continuaron su etapa de engorda. En aquellos sitios que continuaron creciendo las pérdidas de peces fueron: Mín. 15 %, Máx. 65 %, y Media 45 %. Los datos fueron recopilados por la autoridad sanitaria nacional competente a principios de junio de 2021, de 150 granjas. 87 de estos fueron identificados como casos en ese momento. Los datos disponibles de la encuesta incluyen:

- SiteName (Nombre centro)
- ManagementArea (Area de Manejo)
- Case (Caso 1/0)
- MeanWaterTemperature (Temperatura promedio)
- Density (Densidad)
- Company (Compañía)
- Vaccine (Tipo de vacuna)
- SeaLice (Presencia de Sealice)
- GillDisease (Enfermedad branquial)
- FailedSmolt (Calidad smolt)

Descripción de las variables en estudio, factores a analizar y numero total de observaciones.

Se realizara un analisis exploratorio a la base de datos de 150 centros de cultivo, con su correspondientes informacion

```
D <- read_delim("MarineFjordland.txt")
```

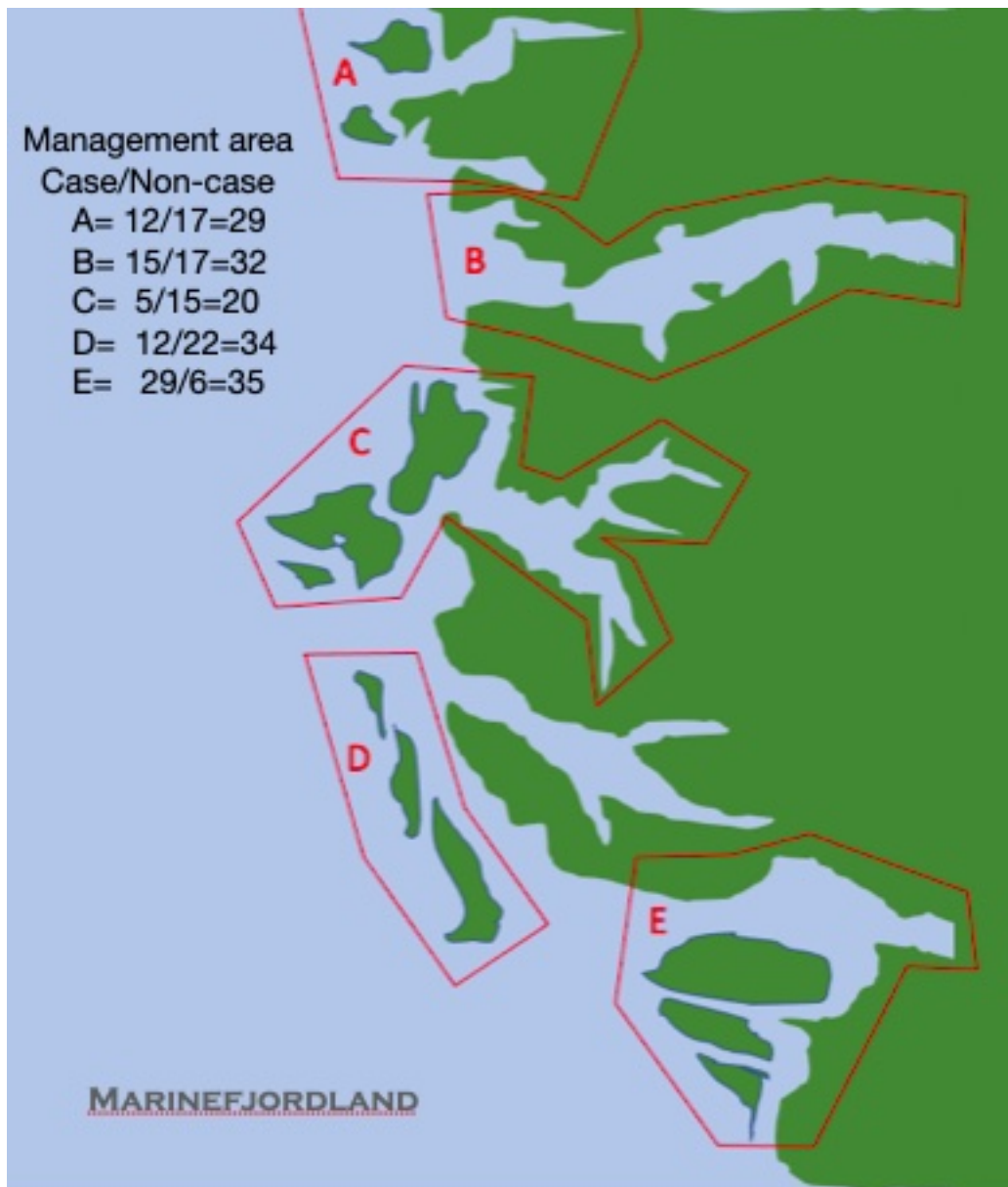


Figure 1: Figura 1: Areas de Manejo A, B, C, D y E de la zona MarineFjordland

```
## New names:
## Rows: 150 Columns: 11
## -- Column specification
## ----- Delimiter: "\t" chr
## (4): SiteName, ManagementArea, Company, Vaccine dbl (6): Case,
## MeanWaterTemperature, Density, SeaLice, GillDisease, FailedSmolt lgl (1): ...11
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `` -> `...11`
```

Preguntas Iniciales

¿Cuántas y qué tipo de variables se dispone para el análisis? 8 variables: Area de manejo (cualitativa nominal), Compañía (cualitativa nominal), Tipo de vacuna (cualitativa nominal), Temperatura promedio (cuantitativo continuo), Densidad (cuantitativo continuo), Presencia de Sealice (variable aleatoria discreta con distribucion bernoulli), enfermedad branquial (variable aleatoria discreta con distribucion bernoulli), calidad de smolt (variable aleatoria discreta con distribucion bernoulli). ¿Cuales son los tratamientos? Caso 1 (positivo, es decir mortalidad mayor al 5% diario), Caso 0 (negativo, es decir no se mortalidad mayor al 5% diario), variable aleatoria discreta con distribucion bernoulli ¿La base de datos está completa?, ¿tiene errores? ¿Es posible responder las causas de mortalidad con los datos disponibles? ¿la cantidad de datos y variables permite hacer un análisis estadístico?

```
table(D$Case,D$SiteName )
```

```
##
##      SiteName1 SiteName10 SiteName100 SiteName101 SiteName102 SiteName103
## 0           1           1           1           0           1           1
## 1           0           0           0           1           0           0
##
##      SiteName104 SiteName105 SiteName106 SiteName107 SiteName108 SiteName109
## 0           1           0           0           0           1           1
## 1           0           1           1           1           0           0
##
##      SiteName111 SiteName110 SiteName111 SiteName112 SiteName113 SiteName114
## 0           1           0           1           0           1           1
## 1           0           1           0           1           0           0
##
##      SiteName115 SiteName116 SiteName117 SiteName118 SiteName119 SiteName12
## 0           1           1           0           0           1           1
## 1           0           0           1           1           0           0
##
##      SiteName120 SiteName121 SiteName122 SiteName123 SiteName124 SiteName125
## 0           0           1           1           0           0           1
## 1           1           0           0           1           1           0
##
##      SiteName126 SiteName127 SiteName128 SiteName129 SiteName13 SiteName130
## 0           0           0           1           0           0           0
## 1           1           1           0           1           1           1
##
##      SiteName131 SiteName132 SiteName133 SiteName134 SiteName135 SiteName136
## 0           0           0           0           0           1           0
## 1           1           1           1           1           0           1
##
##      SiteName137 SiteName138 SiteName139 SiteName14 SiteName140 SiteName141
```

```

## 0      0      1      1      0      0      1
## 1      1      0      0      1      1      0
##
##      SiteName142 SiteName143 SiteName144 SiteName145 SiteName146 SiteName147
## 0      1      1      1      1      1      0
## 1      0      0      0      0      0      1
##
##      SiteName148 SiteName149 SiteName15 SiteName150 SiteName16 SiteName17
## 0      0      0      1      1      1      1
## 1      1      1      0      0      0      0
##
##      SiteName18 SiteName19 SiteName2 SiteName20 SiteName21 SiteName22 SiteName23
## 0      1      1      1      1      1      0      1
## 1      0      0      0      0      0      1      0
##
##      SiteName24 SiteName25 SiteName26 SiteName27 SiteName28 SiteName29 SiteName3
## 0      1      1      1      1      0      0      1
## 1      0      0      0      0      1      1      0
##
##      SiteName30 SiteName31 SiteName32 SiteName33 SiteName34 SiteName35
## 0      0      1      0      0      0      1
## 1      1      0      1      1      1      0
##
##      SiteName36 SiteName37 SiteName38 SiteName39 SiteName4 SiteName40 SiteName41
## 0      0      1      0      0      0      1      1
## 1      1      0      1      1      1      0      0
##
##      SiteName42 SiteName43 SiteName44 SiteName45 SiteName46 SiteName47
## 0      0      0      1      1      0      0
## 1      1      1      0      0      1      1
##
##      SiteName48 SiteName49 SiteName5 SiteName50 SiteName51 SiteName52 SiteName53
## 0      1      1      0      0      1      0      1
## 1      0      0      1      1      0      1      0
##
##      SiteName54 SiteName55 SiteName56 SiteName57 SiteName58 SiteName59 SiteName6
## 0      0      0      1      1      1      0      1
## 1      1      1      0      0      0      1      0
##
##      SiteName60 SiteName61 SiteName62 SiteName63 SiteName64 SiteName65
## 0      1      1      1      0      0      0
## 1      0      0      0      1      1      1
##
##      SiteName66 SiteName67 SiteName68 SiteName69 SiteName7 SiteName70 SiteName71
## 0      0      1      1      0      0      1      0
## 1      1      0      0      1      1      0      1
##
##      SiteName72 SiteName73 SiteName74 SiteName75 SiteName76 SiteName77
## 0      0      0      1      0      1      1
## 1      1      1      0      1      0      0
##
##      SiteName78 SiteName79 SiteName8 SiteName80 SiteName81 SiteName82 SiteName83
## 0      0      0      0      1      1      1      1
## 1      1      1      1      0      0      0      0

```

```
##
##      SiteName84 SiteName85 SiteName86 SiteName87 SiteName88 SiteName89 SiteName9
##      0          1          0          1          0          0          1          0
##      1          0          1          0          1          1          0          1
##
##      SiteName90 SiteName91 SiteName92 SiteName93 SiteName94 SiteName95
##      0          0          0          0          0          0          1
##      1          1          1          1          1          1          0
##
##      SiteName96 SiteName97 SiteName98 SiteName99
##      0          0          1          0          0
##      1          1          0          1          1
```

Resumen y Visualización

```
summary(D)
```

```
##      SiteName      ManagementArea      Case      MeanWaterTemperature
## Length:150      Length:150      Min.      :0.0000      Min.      : 6.100
## Class :character Class :character 1st Qu.:0.0000      1st Qu.: 8.025
## Mode  :character Mode  :character Median :0.0000      Median : 9.050
##                                     Mean  :0.4867      Mean   : 9.052
##                                     3rd Qu.:1.0000      3rd Qu.:10.475
##                                     Max.   :1.0000      Max.   :11.800
##      Density      Company      Vaccine      SeaLice
## Min.      : 3.200      Length:150      Length:150      Min.      :0.0000
## 1st Qu.: 8.457      Class :character Class :character 1st Qu.:0.0000
## Median :10.800      Mode  :character Mode  :character Median :1.0000
## Mean      :12.268                                     Mean   :0.6333
## 3rd Qu.:14.400                                     3rd Qu.:1.0000
## Max.      :32.800                                     Max.   :1.0000
##      GillDisease      FailedSmolt      ...11
## Min.      :0.0      Min.      :0.0000      Mode:logical
## 1st Qu.:0.0      1st Qu.:0.0000      NA's:150
## Median :0.0      Median :0.0000
## Mean      :0.3      Mean      :0.2533
## 3rd Qu.:1.0      3rd Qu.:0.7500
## Max.      :1.0      Max.      :1.0000
```

Centros de Cultivo por area de Manejo (Barrio)

```
table(D$ManagementArea)
```

```
##
##  A  B  C  D  E
## 29 32 20 34 35
```

Casos Totales 0=negativo 1= caso positivo

```
table(D$Case, D$ManagementArea)
```

```
##
```

```
##      A  B  C  D  E
##    0 17 17 15 22  6
##    1 12 15  5 12 29
```

Histograma Temperatura promedio

```
ggplot(D, aes(x=D$MeanWaterTemperature))+
  geom_histogram(color="green", fill="lightblue", bins = 10)+theme_classic()+theme(text = element_text(
```

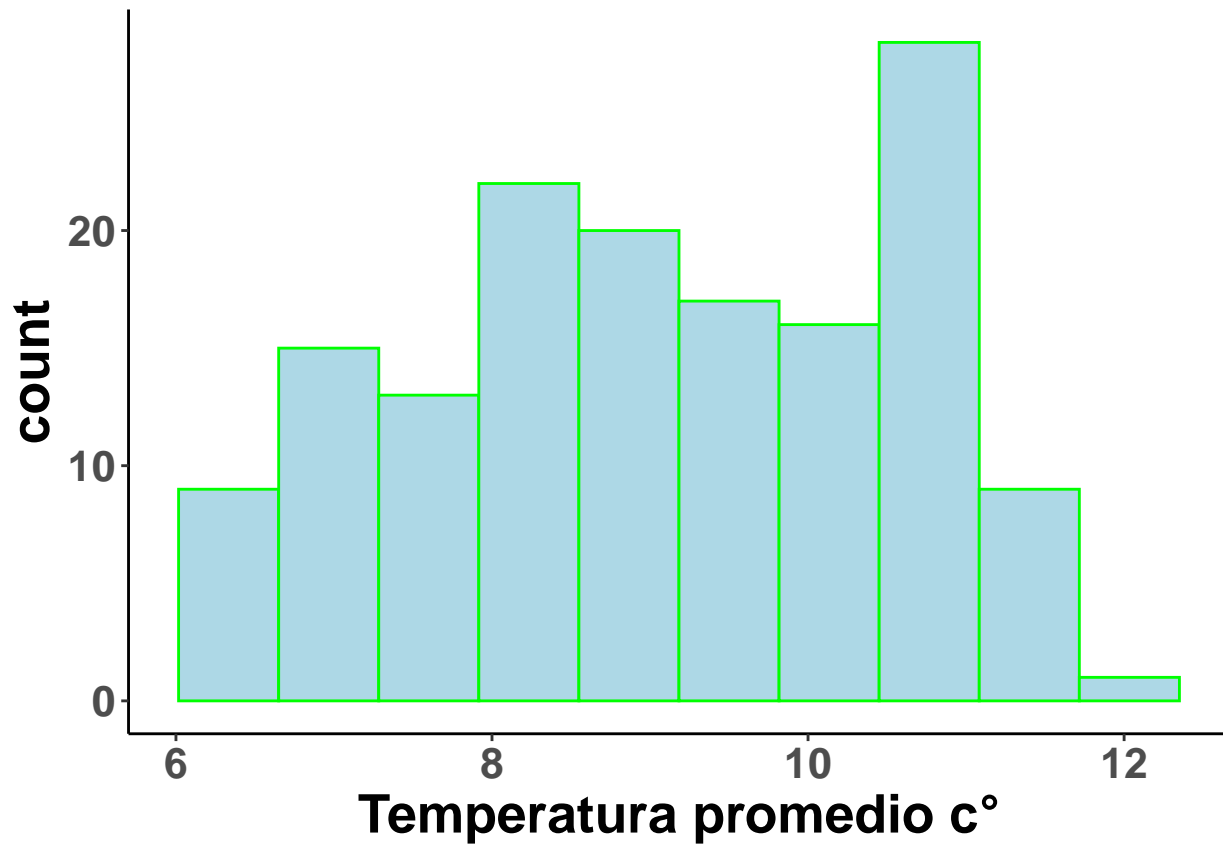
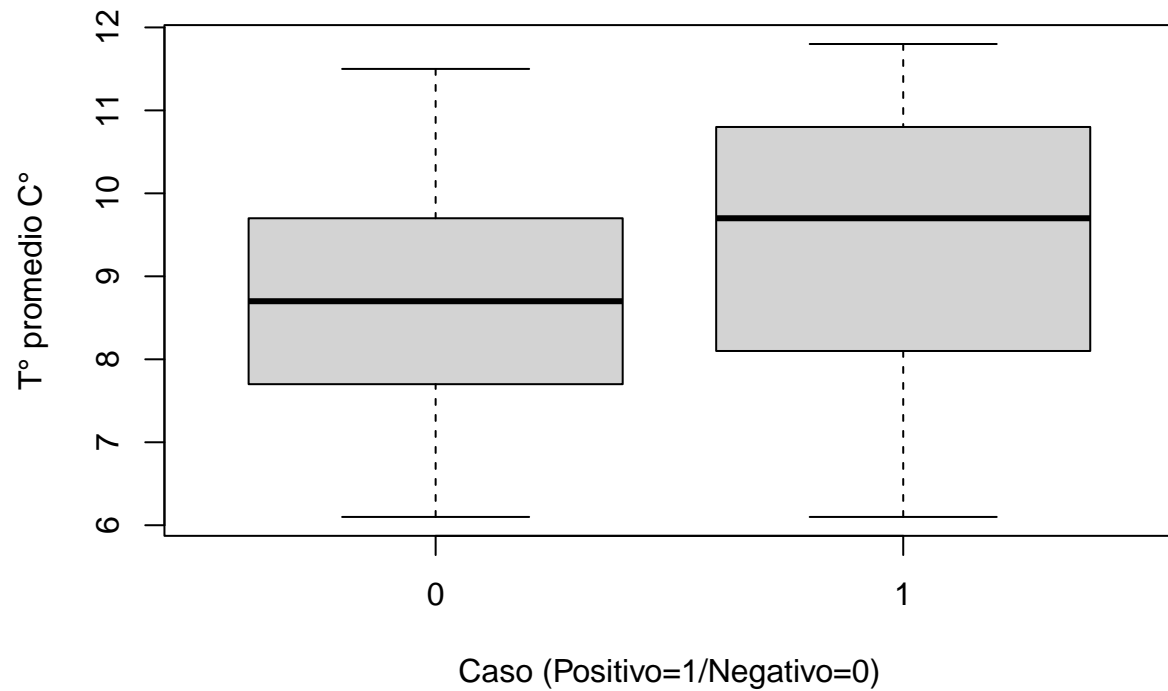


Gráfico Bloxpot asociando T° promedio con Caso

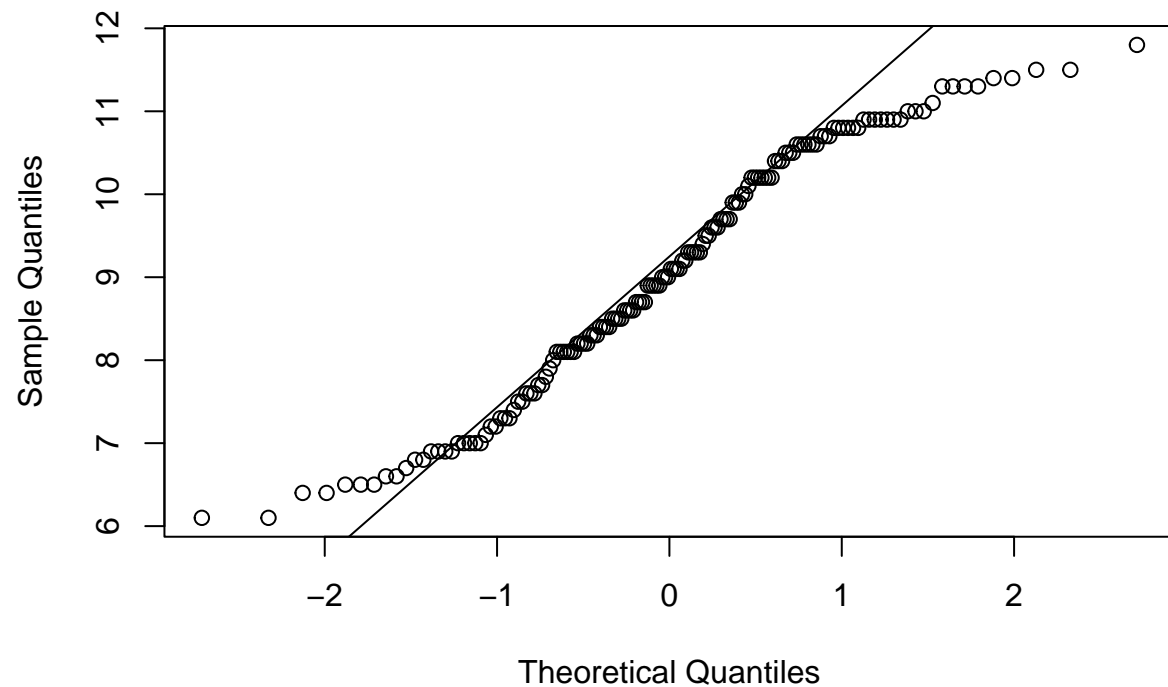
```
boxplot(D$MeanWaterTemperature~D$Case, xlab= "Caso (Positivo=1/Negativo=0)", ylab="T° promedio C°")
```



QQplot para ver la normalidad de los datos

```
qqnorm(D$MeanWaterTemperature)
qqline(D$MeanWaterTemperature)
```

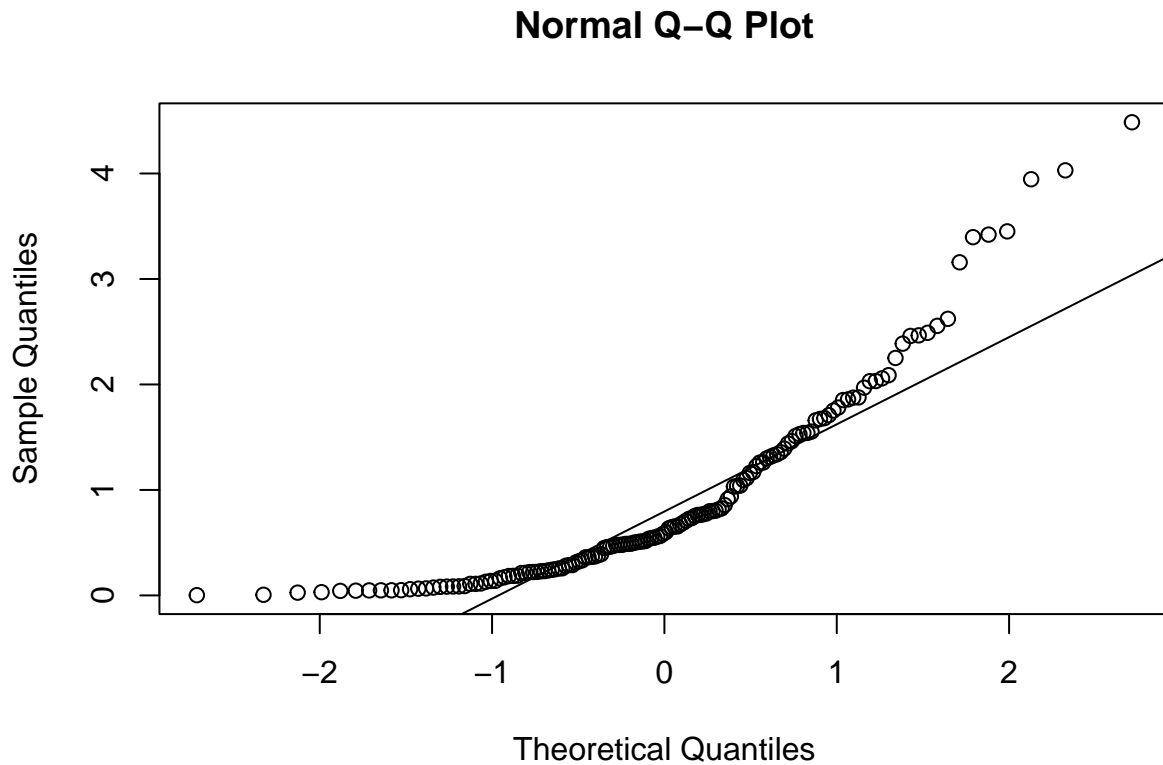
Normal Q-Q Plot



En este caso, buscaremos la normalidad con otra prueba, ya que la gran mayoría de las observaciones se

encuentran en el eje, por lo que la interpretación del valor de P, debería ayudarnos. Otra opción sería realizar un gráfico con una distribución gamma confirmando que los datos no tienen una distribución normal y una asimetría a la izquierda.

```
D$MeanWaterTemperature <- rgamma (150, 1)
qqnorm (D$MeanWaterTemperature)
qqline (D$MeanWaterTemperature)
```



Histograma Densidad (distribucion binomial con desplazamiento a la derecha)

```
ggplot(D, aes(x=D$Density))+
  geom_histogram(color="green", fill="lightblue", bins = 10)+theme_classic()+theme(text = element_text(
```

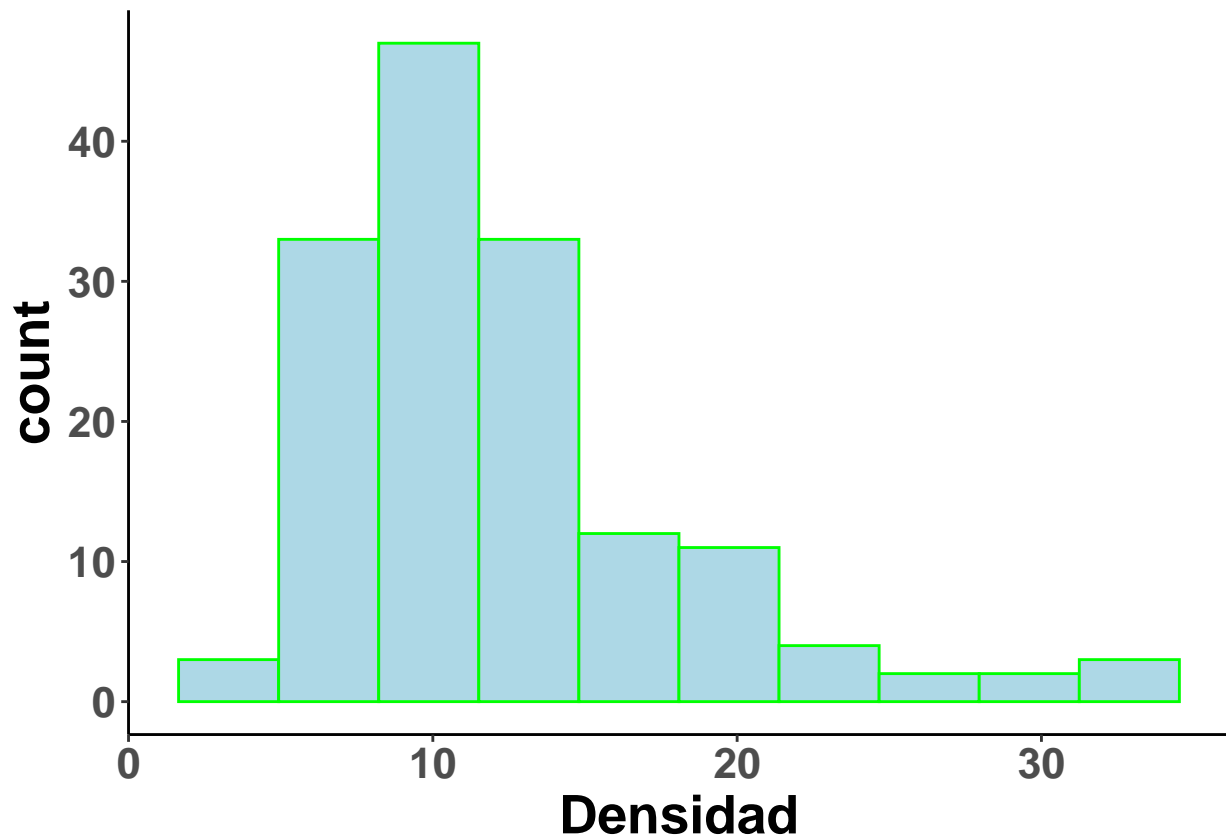
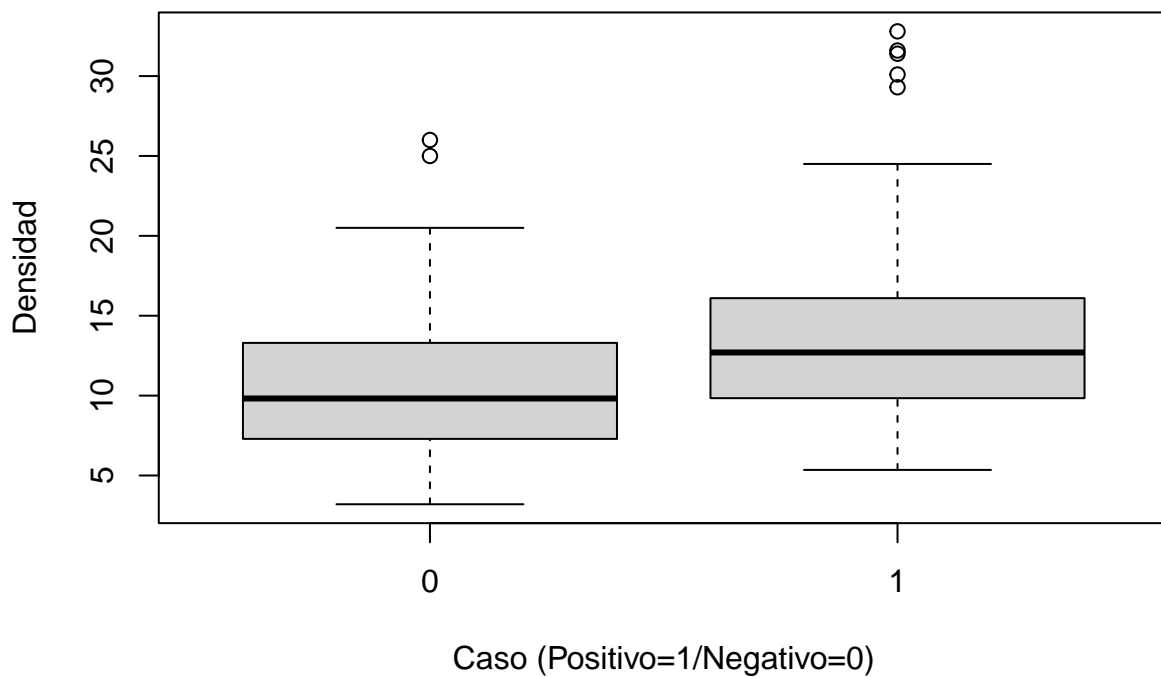



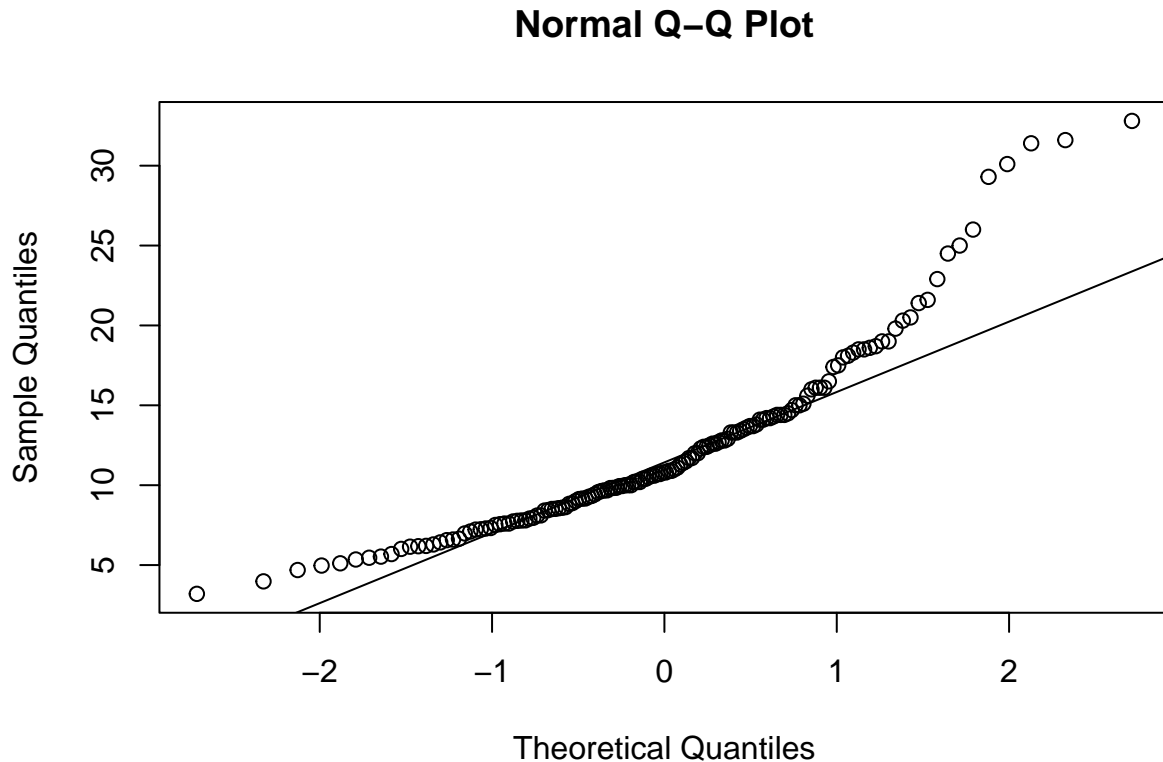
Gráfico Bloxpot asociando Densidad con Caso

```
boxplot(D$Density~D$Case, xlab= "Caso (Positivo=1/Negativo=0)", ylab="Densidad")
```



QQplot para ver la normalidad de los datos

```
qqnorm(D$Density)  
qqline(D$Density)
```



De acuerdo a lo que vemos en el qqplot, los datos de Densidad no tienen una distribución normal y tienen una asimetría hacia la izquierda.

Preguntas Adicionales: Hipotesis nula

¿La mortalidad observada en los casos positivos está explicada por tipo de vacuna, zona de manejo, temperatura promedio del agua, densidad de cultivo, compañía productora, enfermedad de branquias, presencia de caligus, calidad de smolt?

Comentario Final

La base de datos está limpia, completa, existen variables cualitativas y cuantitativas. Las variables cuantitativas no tienen una distribución normal por ende se requiere pruebas no paramétricas para un adecuado análisis estadístico.