

Presentacion Final Diplomado Analisis de Datos con R

Juan José Rodríguez Maulén

2022-06-30

Analisis comparativo de Proveedor de Fotoperiodo y Centros de Cultivo de Salmon Coho sobre la Maduración

- 1). Descripción del Problema**
- 1). Descripción de las variables de estudio**
- 3). Métodos Estadísticos**
- 4). Exploración de Datos (EDA)**
- 5). Hipótesis**
- 6). Resultados**
- 4). Conclusiones**

Descripción del Problema

Se observó madurez en peces cultivados en la región de Los Lagos, sometidos a regimen de fotoperiodo para prevenir la madurez. El proveedor de fotoperiodo correspondió a la empresa BIOLED quienes utilizaron 3 intensidades lumínicas (W) en centros de cultivos de peces provenientes de las pisciculturas Huincara, Coipue, Lican y Lago Rupanco. En terminos de madurez observada por mix de jaulas, estas fluctuaron entre un 0 a un 20%, mientras que a nivel de centro de cultivo, esta alcanzo un 6,81% siendo aceptable un 5%.

FACTORES A ANALIZAR Y NUMERO DE OBSERVACIONES

Los datos de madurez, correspondieron a las observaciones realizadas en plantas de proceso, para la clasificación de calidades, donde una de las causas de degradación correspondio a madurez por jaula y centro de cultivo

Variable respuesta= % de maduración (Cuantitativa Discreta)

Variable explicativa= Proveedor de Fotoperiodo y Centro de Cultivo (Cualitativa Nominal) N= 220 observaciones.

Para el presente reporte final, la estrategia correspondio al análisis de las categorías Proveedor de Fotoperiodo y Centro de Cultivo (Análisis Univariado de Dos Vías)

Se realizó un análisis exploratorio de datos, para observar patrones, posteriormente se realizó la evaluación de supuestos a los residuales de los datos, con la finalidad de determinar el método estadístico a utilizar.

Dados los resultados gráficos y los p-value resultantes de los tests estadísticos para la evaluación de supuestos, se demostró que los residuales no provenían de una distribución normal. De acuerdo a lo anterior, se ajustaron Modelos Lineales Generalizados y se seleccionó el mejor modelo a través del Criterio de Akaike (AIC).

Como análisis a posteriori, se realizó un test de Tukey para determinar que grupos eran diferentes entre sí.

Exploración de Datos (EDA)

Evaluación balanceo de datos **Proveedor**

Para el caso de la categoría Proveedor, se puede observar que el nivel “Bioled” tiene mayor número de observaciones, por lo tanto, no está totalmente balanceado.

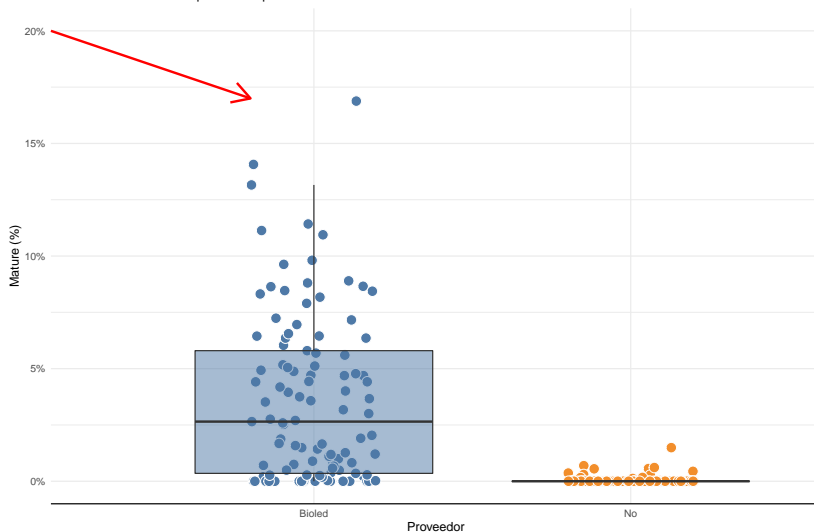
```
##
```

```
## Bioled      No
```

```
##      60      19
```

Boxplot para categoría Proveedor

% Madurez en planta de proceso según Fotoperiodo
Dato a nivel de unidad de proceso en planta

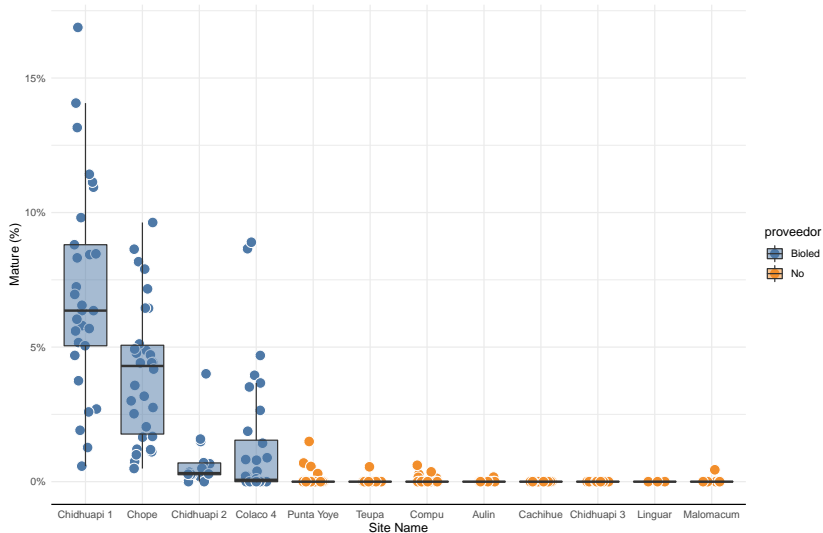


Por otra parte, se observaron datos atípicos (e.g., Flecha Roja)

Boxplot para categoría Proveedor*Centro

% Madurez en planta segun Centro y Fotoperiodo

Dato a nivel de unidad de proceso en planta



Hipotesis por Proveedor de Fotoperiodo & Centro de Cultivo

Hipótesis Biológica = Centros con fotoperiodo, no presentan razgos de madurez

Hipótesis Estadística = Centros con fotoperiodo no presentan diferencias significativas en el desarrollo de madurez respecto a centros sin fotoperiodo.

Ajuste de Modelo

Variable Respuesta = % de maduración (Cuantitativa Discreta)

Variable Explicativa = Categoría Proveedor (Bioled & Control) *
Centro de Cultivo

$$y \sim B0 + BX1 + BX2 + BX1 * BX2 + E$$

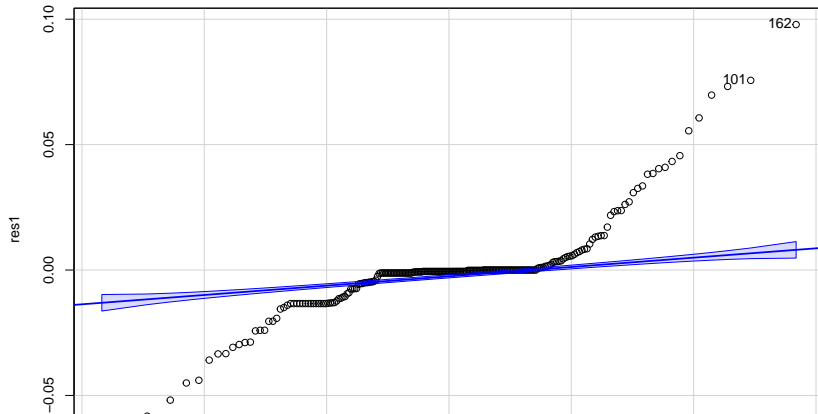
Evaluación de Supuestos

```
## [1] "Jaula individual" "Centro" "Cage"
## [5] "Maduro" "% Desadaptado" "Origen"
## [9] "Fotoperiodo" "Proveedor" "Potencia"

## Jaula individual Centro Cage
## Min. :0.0000 Length:220 Length:220
## 1st Qu.:0.0000 Class :character Class :character
## Median :0.0000 Mode :character Mode :character
## Mean :0.3636
## 3rd Qu.:1.0000
## Max. :1.0000
##
## Maduro % Desadaptado Origen
## Min. :0.00000 Min. :0.00000 Length:220
## 1st Qu.:0.00000 1st Qu.:0.00355 Class :character
## Median :0.00000 Median :0.01276 Mode :character
## Mean :0.01760 Mean :0.02100
## 3rd Qu.:0.00161 3rd Qu.:0.00070
```

Evaluación de Normalidad

```
##  
## Shapiro-Wilk normality test  
##  
## data:  res1  
## W = 0.78278, p-value < 2.2e-16
```



Selecciono el mejor modelo a través de AIC

```
##          df          AIC
## lm1    13 -1090.64823
## glm2   12           Inf
## glm3   13    53.98772
## glm4   12    31.95718
```

```
## Analisis de Varianca de Dos Vias lm1
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: Maduro
```

```
##          Df    Sum Sq  Mean Sq F value    Pr(>F)
## Proveedor    1 0.070471 0.070471 182.164 < 2.2e-16 ***
## Centro       10 0.063824 0.006382  16.498 < 2.2e-16 ***
## Residuals   208 0.080466 0.000387
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
```

Test a Posteriori

Aulin Cachihue Chidhuapi 1 Chidhuapi 2 Chidhuapi 3 Chope Colaco
4 "a" "a" "c" "a" "a" "b" "a" Compu Linguar Malomacum Punta
Yoye Teupa "a" "a" "a" "a" "a"

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-  
## (as 'lib' is unspecified)
```

```
## Loading required package: mvtnorm
```

```
## Loading required package: survival
```

```
## Loading required package: TH.data
```

```
##
```

```
## Attaching package: 'TH.data'
```

```
## The following object is masked from 'package:MASS':
```

```
##
```

```
##      geyser
```

Conclusiones

De acuerdo a la exploración de datos, se puede evidenciar que el proveedor fotoperiodo, presento maduración en los centros de cultivos donde se implentó la estrategia en comparación con los centros de cultivo donde no se encontraba implementada la estrategia de Fotoperiodo.

El Centro de Cultivo que presentó mayor madurez ($> 5\%$) correspondio al centro Chidhuapi 1

Los residuales de los datos no presentaron una distribucion normal ni homocedasticidad

Se intentaron ajustar Modelos Lineales Generalizados, pero las seleccion por el Criterio de Akaike, no fue coherente, dado que el mejor modelo correspondio al modelo lineal, el cual violo los supuestos.

Aun asi, se realizo el test de Analisis de Varianza, donde se evidencio evidencia estadistica para rechazar la hipotesis de