

Evaluacion de la mortalidad de dos tratamientos en centros de mar y su interacción.

Gustavo Bustos Yáñez

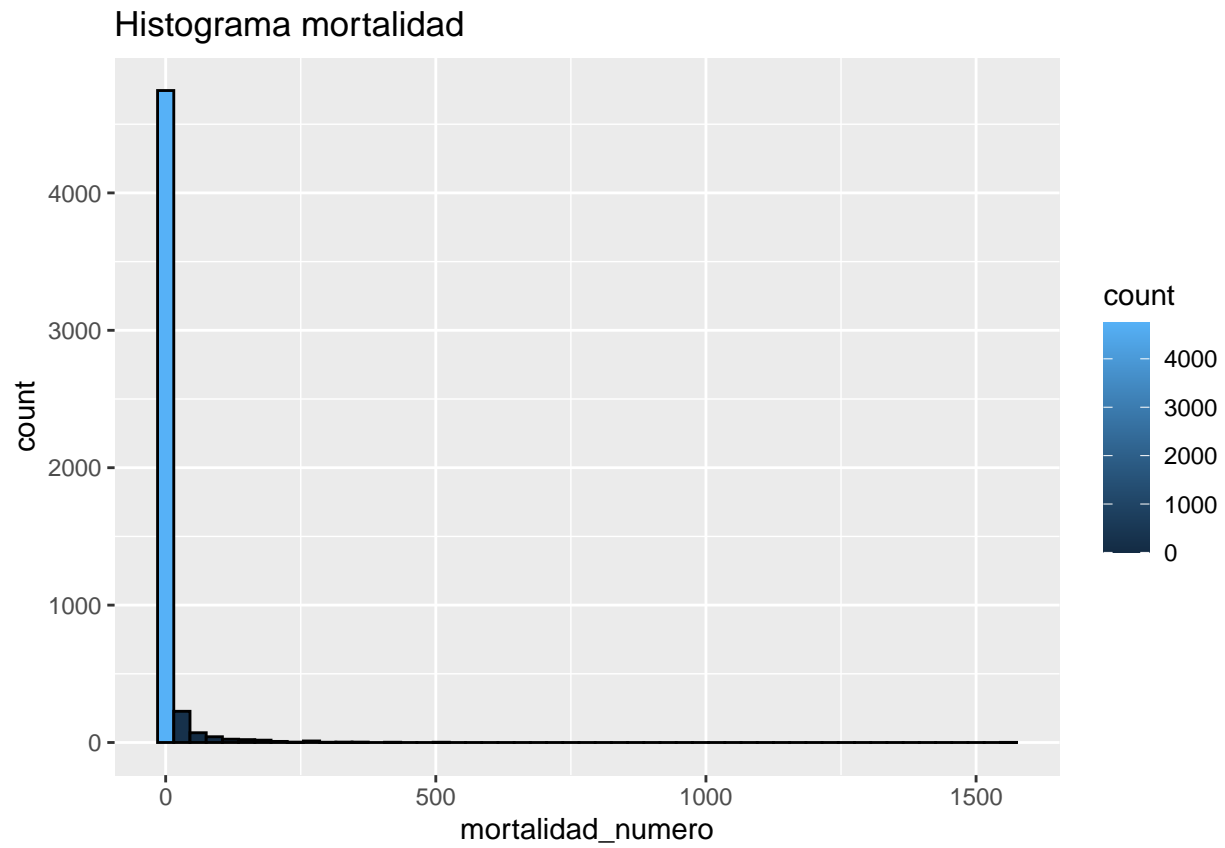
29 November 2021

Exploración de datos.

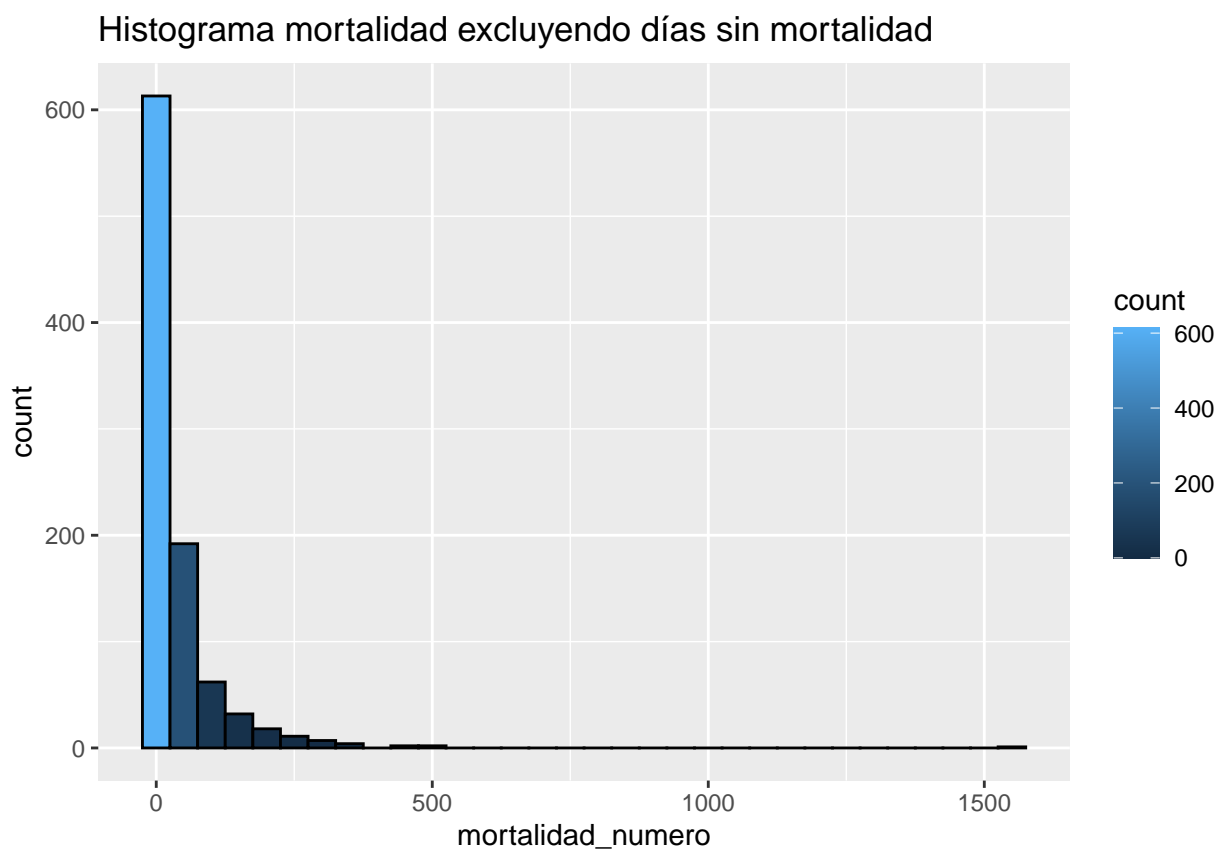
Estructura del set de datos.

```
## tibble [5,182 x 9] (S3: tbl_df/tbl/data.frame)
## $ fecha          : Factor w/ 489 levels "01-01-2020","01-01-2021",...: 176 193 193 209 209 225
## $ unidad         : Factor w/ 14 levels "101","102","103",...: 3 4 3 4 3 4 3 4 3 4 ...
## $ clase          : Factor w/ 4 levels "AA","AB","B",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ tratamiento    : Factor w/ 2 levels "No","Si": 2 2 2 2 2 2 2 2 2 2 ...
## $ vacuna         : Factor w/ 2 levels "No","Si": 1 1 1 1 1 1 1 1 1 1 ...
## $ densidad       : Factor w/ 242 levels "0.2","0.3","0.4",...: 3 3 8 3 8 3 9 7 9 7 ...
## $ origen         : chr [1:5182] "P1" "P1" "P1" "P1" ...
## $ mortalidad_numero : num [1:5182] 0 0 0 0 0 0 0 0 0 0 ...
## $ mortalidad_porcentaje: num [1:5182] 0 0 0 0 0 0 0 0 0 0 ...
```

Histograma



Para ayudar a la visualización del histograma se eliminan los datos en 0, que representan días sin extracción de mortalidad con características que indican la presencia de la patología en los ejemplares extraídos en la mortalidad.



El histograma demuestra que la distribución de los datos no evidencia que sea del tipo normal.

Tabla de frecuencias:

Table 1: Frecuencias variables Tratamiento y Vacuna.

	Sin Vacuna	Con Vacuna
No	0	1380
Si	2607	1195

Table 2: Frecuencias variables Tratamiento y Vacuna en días con mortalidad.

	Sin Vacuna	Con Vacuna
No	0	318
Si	423	203

Table 3: Frecuencia clase.

Clase	Frecuencia
AA	1469
AB	1138
B	1938

Clase	Frecuencia
BA	637

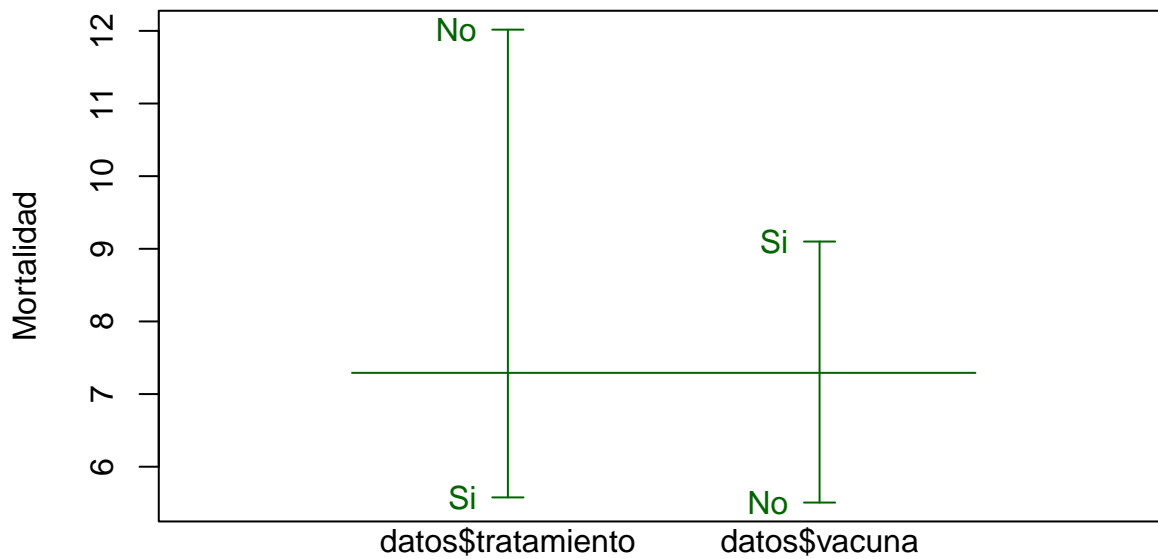
Table 4: Frecuencia según origen de los peces.

Origen	Frecuencia
P1	2607
P2	1938
P3	637

Las tablas de frecuencias demuestran que los datos no se encuentran balanceados, esto puede deberse a que existen unidades con más tiempo de producción que otras y menor cantidad de unidades solamente con el tratamiento. En el caso de clases y Origen se debe a la distinta proporción de peces sembrados en el sitio.

Relación entre variables.

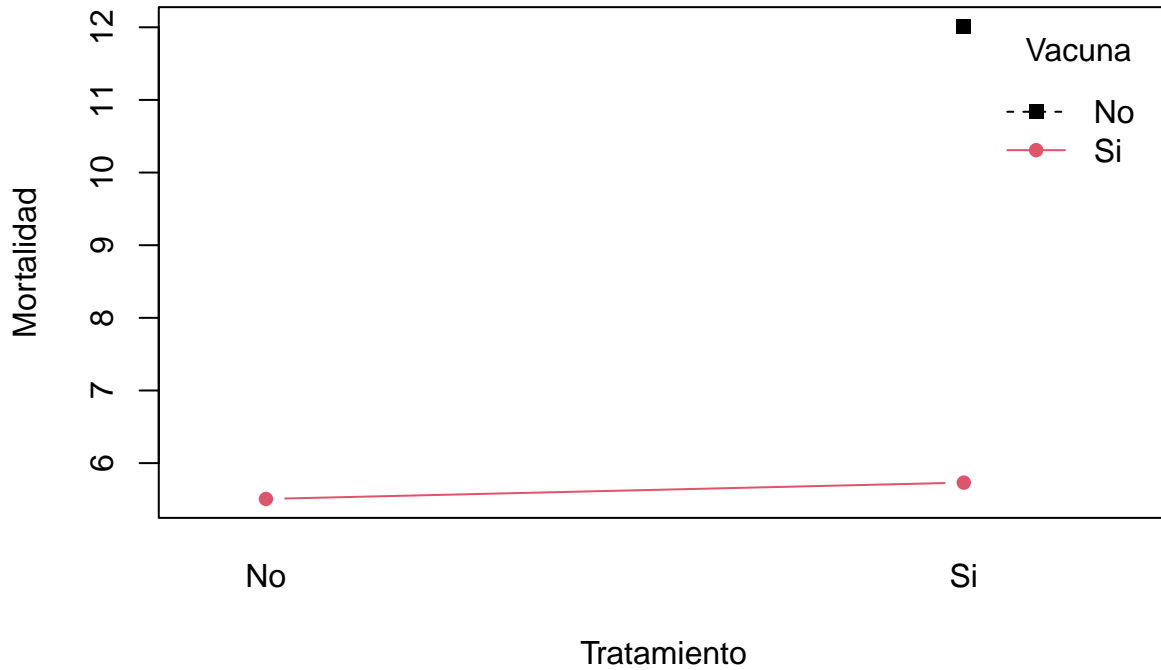
Tamaño de los efectos.



Factores

En la gráfica se puede inferir que el grupo sin tratamiento posee mayor impacto sobre la mortalidad, seguido del grupo sin vacuna.

Gráfico de interacción.



No existe interacción entre las variables tratamiento y vacuna.

Resumen.

Table 5: Resumen variable Tratamiento

tratamiento	varianza	promedio	máximo	mínimo
No	1979.512	12.016667	519	0
Si	1223.888	5.577328	1567	0

Table 6: Resumen variable Vacuna

vacuna	varianza	promedio	máximo	mínimo
No	1568.839	5.506713	1567	0
Si	1289.281	9.099806	519	0

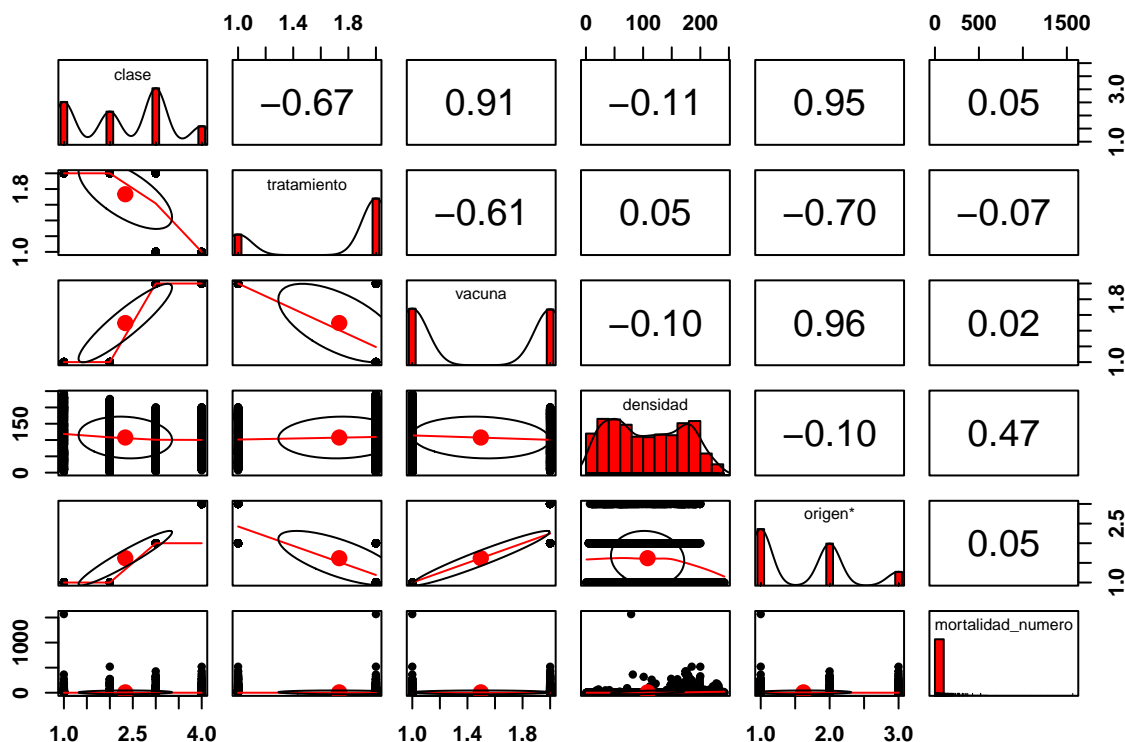
Table 7: Resumen variable Clase

clase	varianza	promedio	máximo	mínimo
AA	2019.7280	4.795779	1567	0
AB	986.5717	6.424429	520	0
B	844.3279	6.754386	427	0
BA	2578.6960	16.235479	519	0

Table 8: Resumen variable Origen

origen	varianza	promedio	máximo	mínimo
P1	1568.8387	5.506713	1567	0
P2	844.3279	6.754386	427	0
P3	2578.6960	16.235479	519	0

En el set de datos existe un gran número de observaciones en 0, lo que representa los días en que no se extrajo mortalidad con características que indiquen la presencia de la patología en los peces extraídos ese día. Existe además, un valor máximo de 1567 que se escapa de los valores registrados regularmente en los días de cultivo, que puede ser considerado como dato atípico.



El factor densidad es el que presenta una mayor correlación con respecto a la mortalidad, sin embargo, esta no es muy elevada. El resto de variables presentan una baja correlación, en el caso de tratamiento, esta es de tipo negativa.

Análisis estadístico

Modelo lineal general.

```
Quitting from lines 191-194 (Prueba1.Rmd) Error in h(simpleError(msg, call)) : error in evaluating the
argument 'object' in selecting a method for function 'summary': object 'mod.1' not found Calls: ... eval ->
-> summary -> .handleSimpleError -> h
```

Modelo lineal mixto

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
## boundary (singular) fit: see ?isSingular
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: mortalidad_porcentaje ~ tratamiento + vacuna + tratamiento:vacuna +
##      (origen | clase)
##      Data: datos
##
## REML criterion at convergence: -12825.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -0.244 -0.113 -0.113 -0.092  60.682
##
## Random effects:
##      Groups   Name                Variance Std.Dev.  Corr
##      clase    (Intercept)  0.000e+00 0.0000000
##              origenP2     9.583e-07 0.0009789   NaN
##              origenP3     5.650e-05 0.0075165   NaN 0.74
##      Residual                4.898e-03 0.0699883
## Number of obs: 5182, groups:  clase, 4
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   0.011810   0.003459   3.414
## tratamientoSi -0.003875   0.003176  -1.220
## vacunaSi      -0.001396   0.002631  -0.531
##
## Correlation of Fixed Effects:
##              (Intr) trtmnS
## tratamientoS -0.918
## vacunaSi     -0.666  0.501
## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

Comparación de modelos

Table 9: Comparación de modelos sin transformación

	df	AIC
MLG	4	-12847
MLM	10	-12805

Con efecto de linealizar se realiza una transformación en base a logaritmo de la variable respuesta. Los valores en 0 se reemplazan por un valor de 0.000001

Modelo lineal general

Quitting from lines 232-235 (Prueba1.Rmd) Error in h(simpleError(msg, call)) : error in evaluating the argument 'object' in selecting a method for function 'summary': object 'mod.1transformado' not found Calls: ... eval -> -> summary -> .handleSimpleError -> h

Modelo lineal mixto

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
## boundary (singular) fit: see ?isSingular

## Linear mixed model fit by REML ['lmerMod']
## Formula: logmortalidad ~ tratamiento + vacuna + tratamiento:vacuna + (origen |
##      clase)
##      Data: datos
##
## REML criterion at convergence: 22993.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.7113  0.3973  0.4194  0.5609  1.2126
##
## Random effects:
##      Groups      Name      Variance Std.Dev. Corr
##      clase      (Intercept) 0.00000 0.0000
##              origenP2     0.03698 0.1923   NaN
##              origenP3     0.36183 0.6015   NaN 0.71
##      Residual              4.93794 2.2221
## Number of obs: 5182, groups:  clase, 4
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  -1.29565    0.11210  -11.558
## tratamientoSi  0.04917    0.10330   0.476
## vacunaSi      0.30612    0.19925   1.536
##
## Correlation of Fixed Effects:
##              (Intr) trtmnS
## tratamientoS -0.922
## vacunaSi     -0.296  0.230
## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

Comparación de modelos

Table 10: Comparación de modelos posterior a la transformación

	df	AIC
logMLG	4	23014
logMLM	10	23014

Una vez realizada la transformación de los datos mediante logaritmo, se obtiene que la información aportada al modelo lineal mixto (logMLM), mejora un poco el modelo, igualando al modelo lineal general (logMLG). Al comparar los datos sin transformar y luego de realizar el tratamiento mediante logaritmo, se obtiene una mejora del coeficiente AIC.

Table 11: Comparación de modelos previo y posterior a la transformación

	df	AIC
MLG	4	-12847
MLM	10	-12805
logMLG	4	23014
logMLM	10	23014