

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

Gustavo Bustos Yáñez

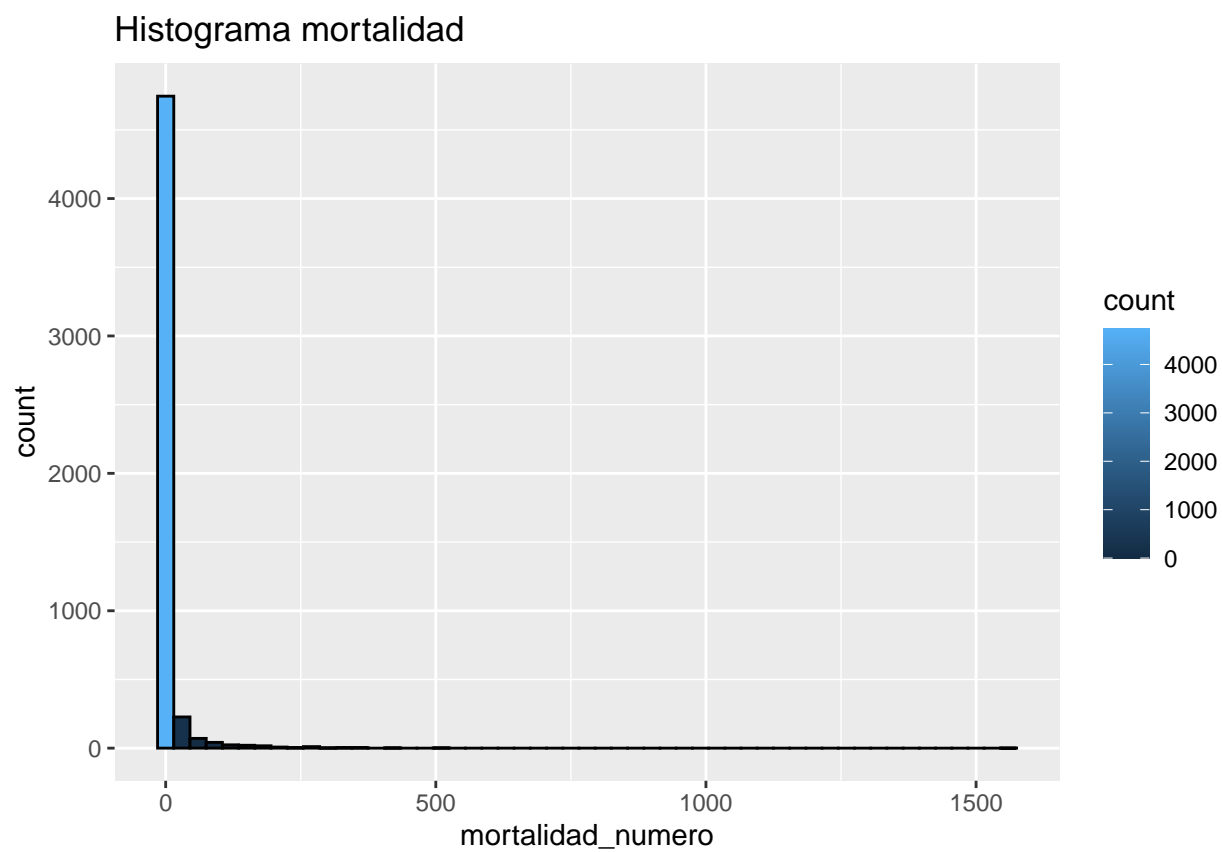
30 November 2021

## *Exploración de datos.*

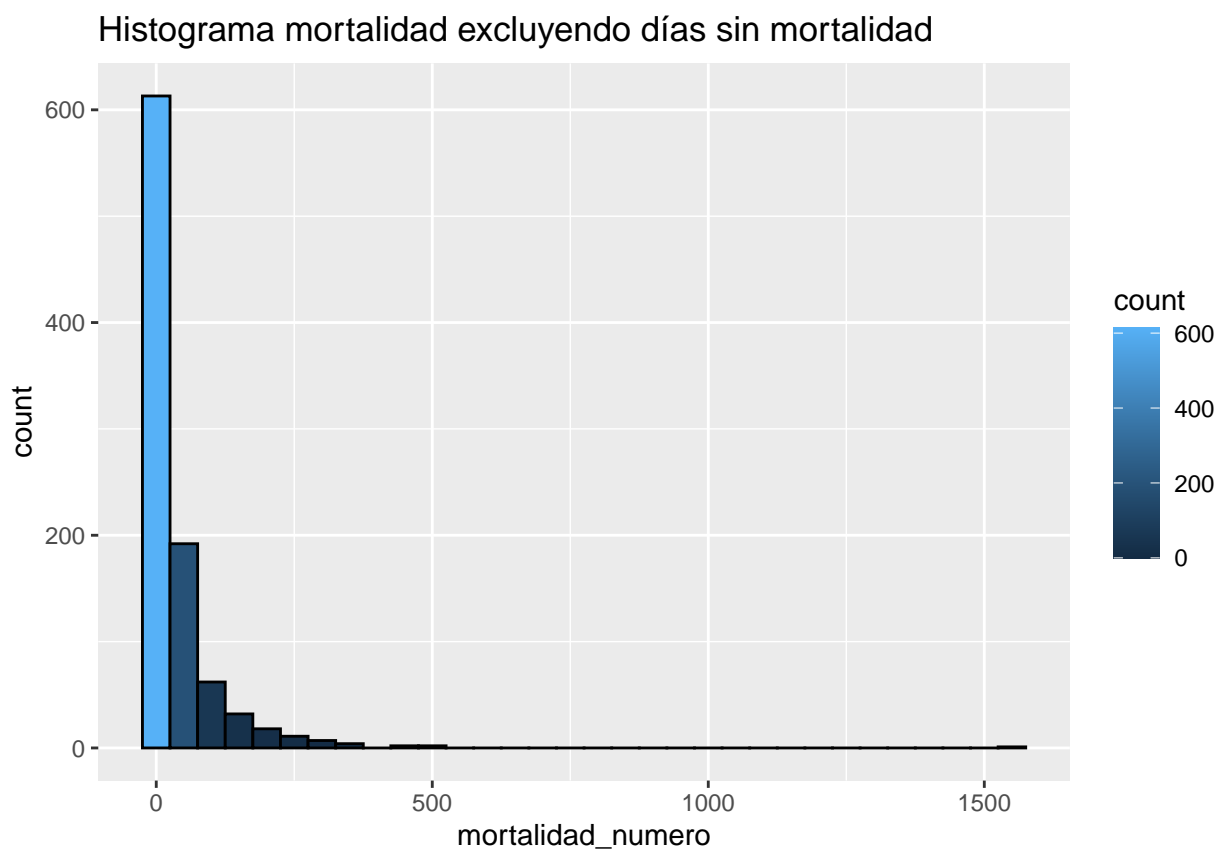
La estructura del set de datos se compone por 5182 observaciones, que corresponden a los días que pasaron las unidades de cultivo en producción. Los datos de mortalidad se encuentran en número total y porcentaje de peces extraídos por unidad el día, los cuales cursan con signología característica de SRS y son clasificados como tal causa.

```
## 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 ...
```

## *Histogramas.*



Para intentar 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 presenta una distribución del tipo normal.

### *Tablas 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 de datos según clase.

Clase	Frecuencia
AA	1469
AB	1138

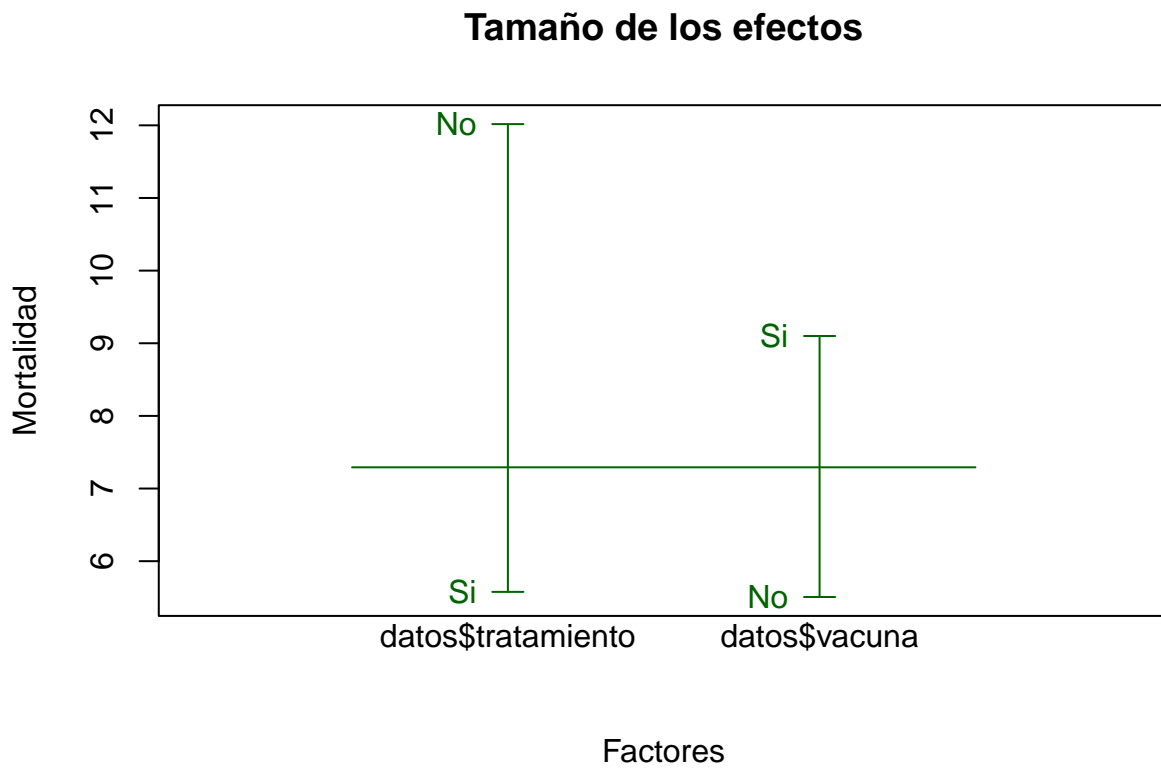
Clase	Frecuencia
B	1938
BA	637

Table 4: Frecuencia según el origen de los peces.

Origen	Frecuencia
P1	2607
P2	1938
P3	637

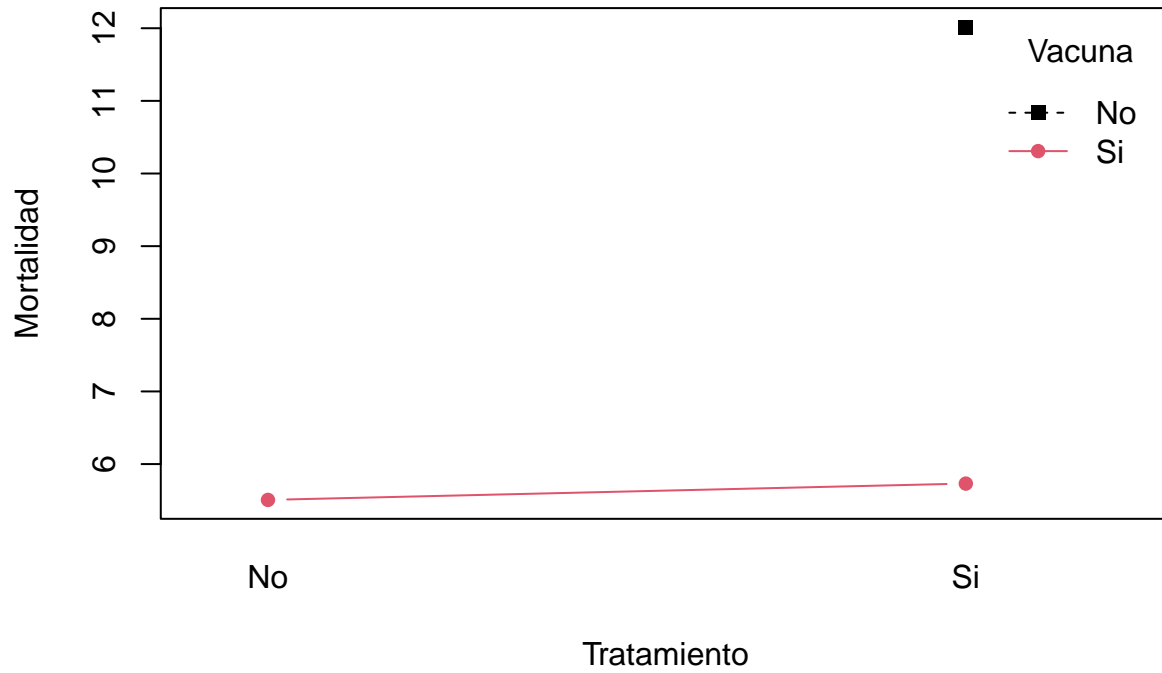
Las tablas de frecuencias demuestran que los datos no se encuentran balanceados. Esto puede deberse a que existen unidades con distinta proporción de peces sembrados con características que presentan desde su origen (piscicultura, tratamiento, vacuna, clase).

### *Relación entre variables.*



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

## Gráfico de interacción



No existe interacción entre las variables tratamiento y vacuna.

## *Resumen set de datos.*

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

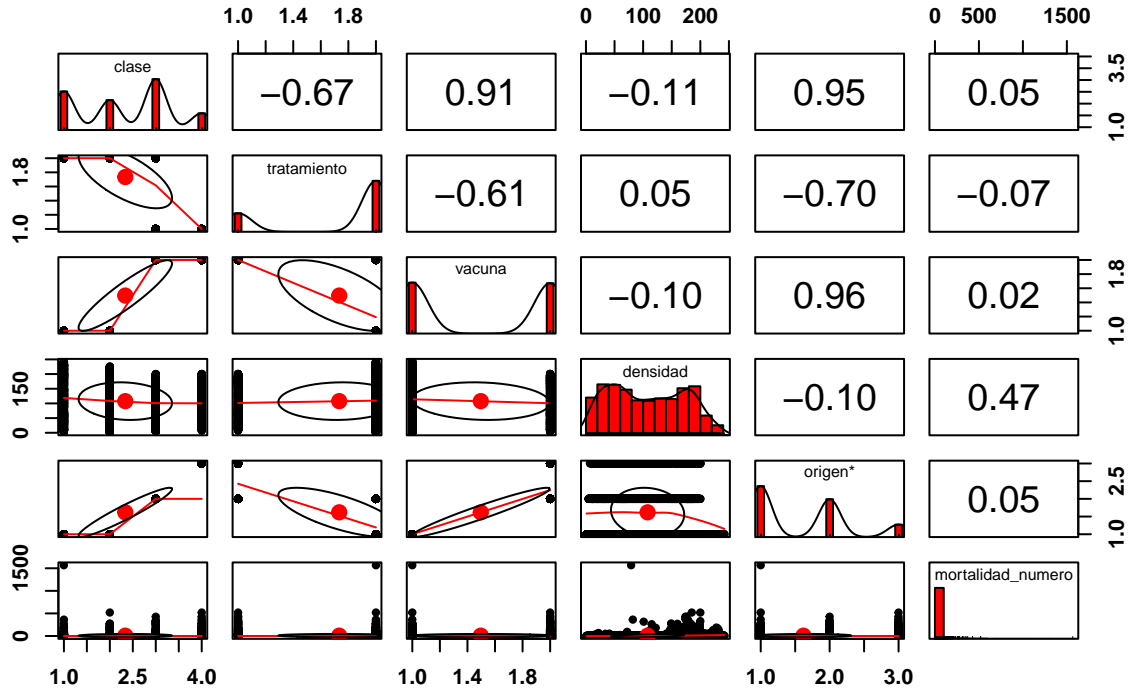
clase	varianza	promedio	máximo	mínimo
AA	2019.7280	4.795779	1567	0
AB	986.5717	6.424429	520	0
B	844.3279	6.754386	427	0
BA	2578.6960	16.235479	519	0

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.

### Correlación entre variables



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.

Para el análisis de regresión lineal se emplearán los datos de porcentaje de mortalidad, producto que el utilizar el número de peces muertos al día genera un error estándar mayor que el emplear el porcentaje.

### Modelo lineal general.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.01494	0.003088	4.839	1.344e-06
tratamientoSi	-0.007006	0.002767	-2.532	0.01136
vacunaSi	-0.00151	0.002446	-0.6172	0.5371

Table 10: Modelo de efectos fijos con interacción

Observations	Residual Std. Error	$R^2$	Adjusted $R^2$
5182	0.07002	0.001493	0.001107

### 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 | densidad)
## Data: datos
##
## REML criterion at convergence: -13060.8
##
## Scaled residuals:
##   Min      1Q  Median      3Q      Max
## -7.117 -0.101 -0.054 -0.026  57.816
##
## Random effects:
##   Groups   Name                Variance Std.Dev. Corr
##   densidad (Intercept) 0.001365  0.03695
##               origenP2   0.001247  0.03531  -0.98
##               origenP3   0.002231  0.04723  -0.80  0.91
##   Residual                0.004294  0.06553
## Number of obs: 5182, groups: densidad, 242
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   0.014550   0.003798   3.831
## tratamientoSi -0.005787   0.002699  -2.144
## vacunaSi      -0.002807   0.003261  -0.861
##
## Correlation of Fixed Effects:
##              (Intr) trtmnS
## tratamientoS -0.690
## vacunaSi     -0.829  0.355
```

```
## 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 11: Comparación de modelos sin transformación

	df	AIC
<b>MLG</b>	4	-12847
<b>MLM</b>	10	-13041

Con efecto de linealizar, se realiza una transformación del porcentaje de mortalidad diario utilizando logaritmo de la variable. Los valores en 0 se reemplazan por un valor de 0.000001 para entregar datos positivos.

## Modelo lineal general con datos transformados.

	Estimate	Std. Error	t value	Pr(> t )
<b>(Intercept)</b>	-1.582	0.09826	-16.1	6.056e-57
<b>tratamientoSi</b>	0.3356	0.08804	3.811	0.0001399
<b>vacunaSi</b>	0.3637	0.07784	4.672	3.058e-06

Table 13: Modelo de efectos fijos con interacción posterior a la transformación de la variable respuesta

Observations	Residual Std. Error	$R^2$	Adjusted $R^2$
5182	2.228	0.004487	0.004103

## Modelo lineal mixto con datos transformados.

```
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
## Linear mixed model fit by REML ['lmerMod']
## Formula: logmortalidad ~ tratamiento + vacuna + tratamiento:vacuna + (origen |
## densidad)
## Data: datos
##
## REML criterion at convergence: 22384.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.4552  0.0238  0.2257  0.5070  1.9289
##
## Random effects:
## Groups   Name                Variance Std.Dev. Corr
## densidad (Intercept) 0.7430   0.8619
##              origenP2  0.9322   0.9655  -0.18
##              origenP3  2.5362   1.5926  -0.48  0.70
```



```
## Residual          3.8804   1.9699
## Number of obs: 5182, groups:  densidad, 242
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  -1.52958    0.10325 -14.814
## tratamientoSi  0.16906    0.08628   1.959
## vacunaSi      0.32199    0.09949   3.236
##
## Correlation of Fixed Effects:
##              (Intr) trtmnS
## tratamientoS -0.764
## vacunaSi     -0.526  0.379
## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

## Comparación de modelos.

Table 14: Comparación de modelos posterior a la transformación

	df	AIC
<b>logMLG</b>	4	23014
<b>logMLM</b>	10	22405

## Conclusión.

Table 15: Comparación de modelos previo y posterior a la transformación

	df	AIC
<b>MLG</b>	4	-12847
<b>MLM</b>	10	-13041
<b>logMLG</b>	4	23014
<b>logMLM</b>	10	22405

Al comparar los datos sin transformar y luego de realizar el tratamiento mediante logaritmo, se obtiene una mejora del coeficiente AIC.

Una vez realizada la transformación de los datos, se obtiene que la información aportada al modelo lineal mixto (logMLM), no mejora el modelo, teniendo un mejor coeficiente AIC el modelo lineal general (logMLG), el cual no incluye los efectos del origen ni densidad.

Con los datos del modelo lineal general una vez transformados (logMLG) podemos inferir que tanto la vacunación como el tratamiento en agua dulce inferen en la mortalidad por causa SRS.