

Los peces y el mercurio
Miguel Alejandro Salas Reyna A00827219

Resumen

En el presente reporte se realiza un análisis estadístico para las estimaciones de mercurio de los lagos de Florida mediante ANOVA y regresión para identificar las variables más significativas en la construcción del modelo. Como veremos más adelante, algunas afectan positivamente este resultado, mientras que otras lo afectan negativamente.

Introducción

El pescado es una fuente importante de nutrientes para el ser humano, siendo uno de los principales alimentos en diversas culturas alrededor del mundo. Sin embargo, altos niveles en la concentración de mercurio en los peces puede ocasionar problemas graves a la salud, llegando incluso a ocasionar envenamamiento por mercurio. Esto es fundamentalmente grave y con repercusiones severas en mujeres embarazadas, que están intentando embarazarse o que están en lactancia para sus bebés.

Debido a esto, se realizó un estudio en 53 lagos de Florida para analizar el nivel de contaminación de mercurio en los peces y cuáles son los factores que más afectan este resultado. Para ello se obtuvieron las siguientes variables:

- X1 = número de indentificación
- X2 = nombre del lago
- X3 = alcalinidad (mg/l de carbonato de calcio)
- X4 = PH
- X5 = calcio (mg/l)
- X6 = clorofila (mg/l)
- X7 = concentración media de mercurio (parte por millón) en el tejido muscular del grupo de peces estudiados en cada lago
- X8 = número de peces estudiados en el lago
- X9 = mínimo de la concentración de mercurio en cada grupo de peces
- X10 = máximo de la concentración de mercurio en cada grupo de peces
- X11 = estimación (mediante regresión) de la concentración de mercurio en el pez de 3 años (o promedio de mercurio cuando la edad no está disponible)
- X12 = indicador de la edad de los peces (0: jóvenes; 1: maduros)

La pregunta es, ¿Cuáles son los principales factores que influyen en el nivel de contaminación por mercurio en los peces de los lagos de Florida? y apartir de aquí surgen diversas dudas y cuestionamientos que se desglozan. El principal objetivo de este documento es resolver algunas de estas cuestiones.

Análisis de los resultados

El primer paso fue obtener datos descriptivos del conjunto de datos. A continuación se muestra la matriz de correlaciones, donde se puede observar que variables como alcalinidad tienen una alta correlación negativa con respecto al mercurio.

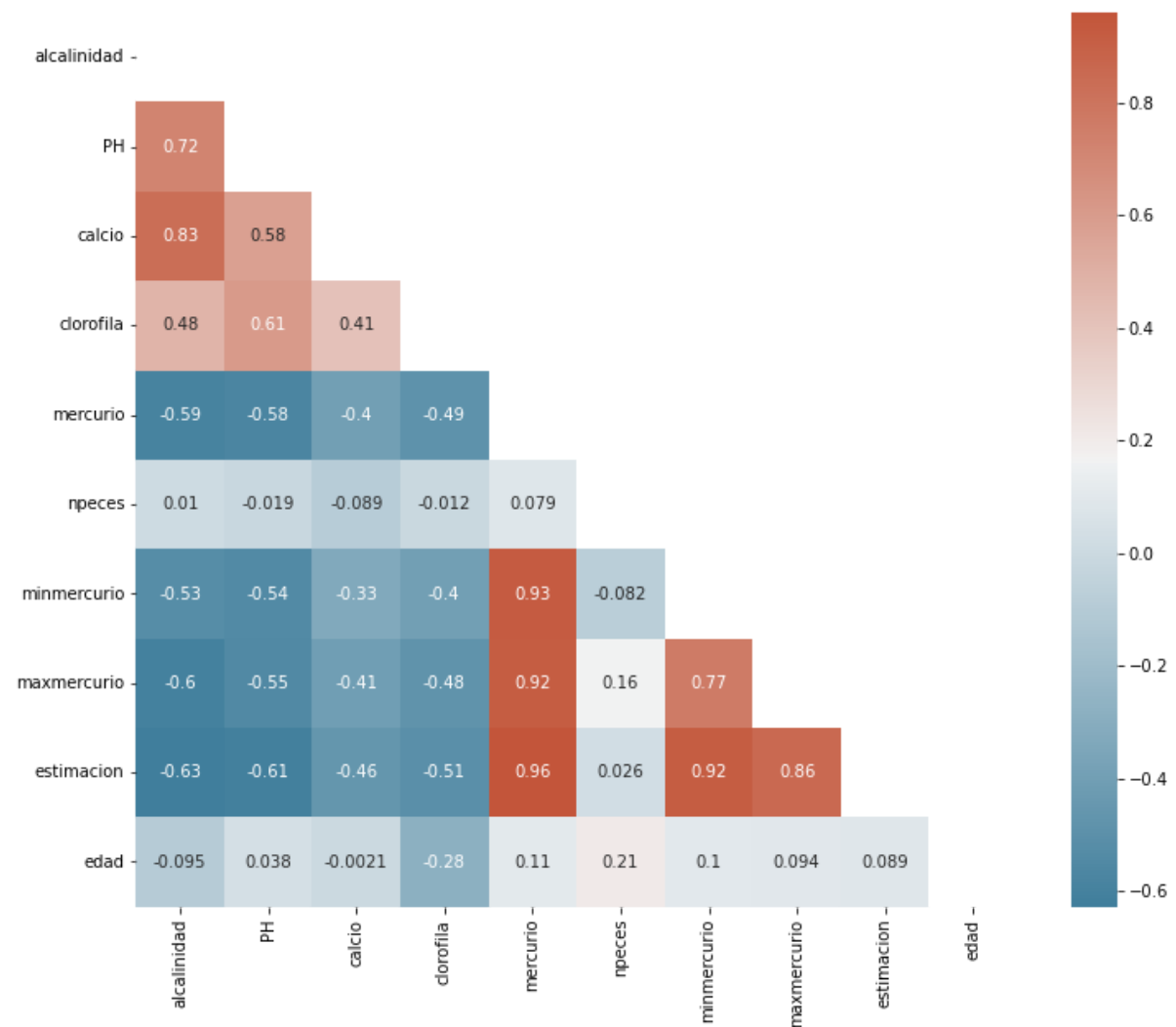
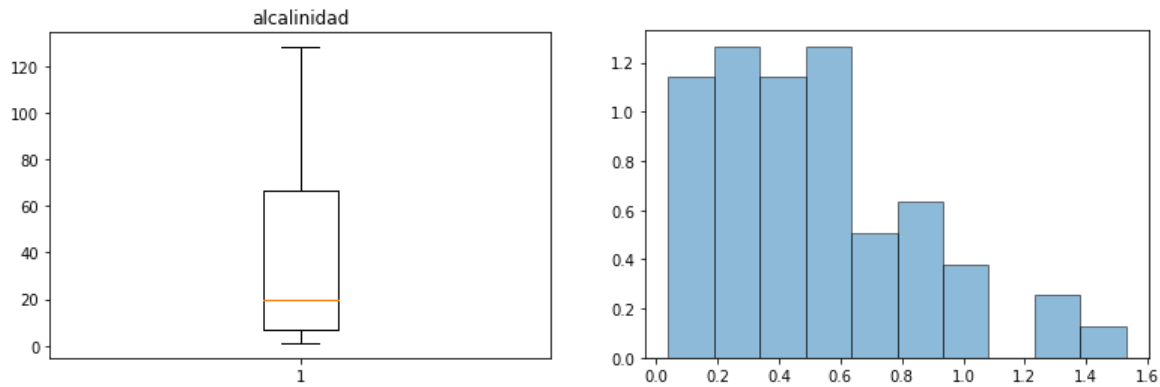
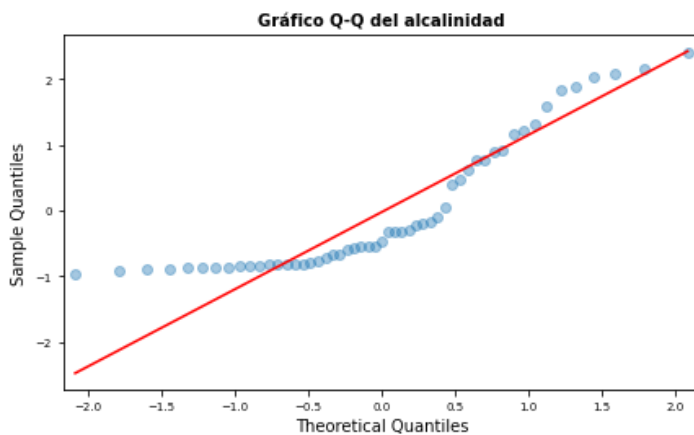


Imagen: Tabla de correlaciones.

Posteriormente se obtuvo boxplots de las variables para identificar valores atípicos y se identificó algunas distribuciones de las variables.



Como se puede observar, la variable de alcalinidad no contiene valores atípicos, y su distribución está un poco cargada a la izquierda. En el posterior análisis de hipótesis se rechazó la normalidad, siendo junto con el calcio las dos únicas variables donde la normalidad se rechazó.



Al final, se realizó una regresión para identificar valores de los coeficientes y pudimos obtener conclusiones interesantes. Para el caso de la regresión, el valor R-squared de 0.46 por lo que se explica el 46% del modelo. Además, algunas variables obtuvieron un p-value menor a 0.05 y otras mayores a este valor. En adición, los coeficientes negativos en algunos casos que afectan negativamente la concentración del mercurio.

Dep. Variable:	mercurio	R-squared:	0.464
Model:	OLS	Adj. R-squared:	0.395
Method:	Least Squares	F-statistic:	6.650
Date:	Sat, 17 Sep 2022	Prob (F-statistic):	4.12e-05
Time:	16:21:54	Log-Likelihood:	-1.1337

No. Observations: 53
AIC: 16.27
Df Residuals: 46
BIC: 30.06
Df Model: 6
Covariance Type: nonrobust

	coef	std err	t	P> t	[0.02 5	0.975]
Intercept	0.9289	0.270	3.444	0.001	0.386	1.472
alcalinidad	-0.0060	0.002	-2.784	0.008	-0.010	-0.002
PH	-0.0391	0.049	-0.798	0.429	-0.138	0.059
calcio	0.0048	0.003	1.728	0.091	-0.001	0.010
clorofila	-0.0026	0.002	-1.564	0.125	-0.006	0.001
npeces	0.0048	0.005	1.054	0.297	-0.004	0.014
edad	-0.0348	0.108	-0.321	0.750	-0.253	0.184

Omnibus: 6.335
Durbin-Watson: 1.468
Prob(Omnibus): 0.042
Jarque-Bera (JB): 5.988
Skew: 0.823
Prob(JB): 0.0501
Kurtosis: 3.042
Cond. No. 518.

Conclusión

Los métodos estadísticos nos permiten obtener información más allá de lo que se nos presenta en los datos y hacer predicciones e inferencias a partir de esta. Para el

caso del mercurio, nos permite hacer predicciones de acuerdo a las variables presentadas y estimar valores futuros de acuerdo a nuevos datos. A partir de esto, se podrá obtener valores del mercurio que nos permita identificar niveles insanos de contaminación en los peces y prevenir riesgos para la salud.

Anexos

<https://github.com/miiguelsr/PortafolioImplementacion.git>