

**Regresión multilineal como apoyo a los análisis proximales**

**Multilinear regression to support proximal analysis**

José Trinidad Ulloa Ibarra<sup>1</sup>, José Trinidad Nieto Navarro<sup>1</sup>, María Inés Ortega Arcga<sup>1</sup>, Juan Felipe Flores Robles<sup>1</sup>, Jaime L. Arrieta Vera<sup>2</sup>

<sup>1</sup> Universidad Autónoma de Nayarit

<sup>2</sup> Universidad Autónoma de Guerrero

Recibido: 19 de febrero de 2019

Aceptado: 27 de mayo de 2019

**Resumen.**

El estudio que se presenta continua con los trabajos de modelación realizados en una comunidad de profesionales de la pesca. Se aborda a la modelación, en este caso multilineal como herramienta aplicada a situaciones propias de la química analítica (análisis proximales), con el objetivo de mostrar que por medio de herramientas de la matemática es posible relacionar los diferentes contenidos químicos del pescado y se llega a concluir que el acercamiento es bastante bueno, pero no lo suficientemente exacto como para plantear la sustitución de los análisis que requieren más tiempo y dinero. La teoría en la que nos basamos, la socioepistemología analiza entre otras cosas el fenómeno educativo considerando el medio en el que se realiza el mismo, por ello sostenemos que realizamos un estudio en una comunidad cuyos resultados al ser transferidos al aula benefician los dos polos en los que se interactúa.

**Palabras clave:** Regresión multilineal, análisis proximales, alimentos, pesca

**Abstract**

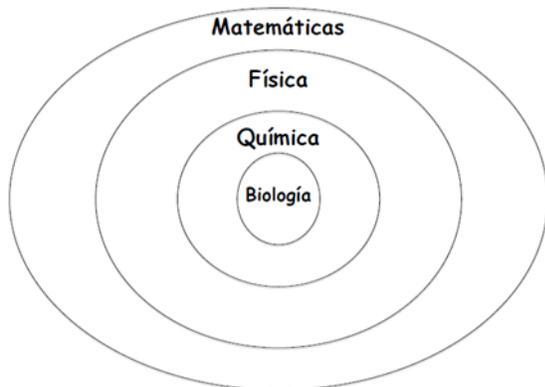
The study that is presented continues with the modeling work carried out in a community of fishing professionals. The modeling is approached, in this case multilinear cone tool applied to situations typical of analytical chemistry (proximal analysis), with the aim of showing that by means of mathematical tools it is possible to relate the different chemical contents of fish and arrive to conclude that the approach is quite good, but not accurate enough to propose the substitution of analyzes that require more time and money. The theory on which we rely, socioepistemology analyzes, among other things, the educational phenomenon considering the medium in which it is carried out, so we maintain that we carry out a study in a community whose results when transferred to the classroom benefit the two poles in which is interacted.

**Key words:** Multilinear regression, proximal analysis, food, fishing

**Introducción.**

Las matemáticas están relacionadas con otras materias y es un base muy importante para ellas. Ciencias como la biología, la fisiología y la medicina en las cuales la matemática no tenía una presencia relevante, están demandando nuevas herramientas matemáticas para poder analizar y explicar muchos problemas sobre los cuales tienen cada vez más información experimental . Como ya se ha citado en otros escritos (Ulloa, Arrieta y Benítez, 2015) la matemática es una herramienta imprescindible para el avance de otras ciencias como la física, la biología, la química, etc

La figura siguiente muestra una concepción de la relación de las matemáticas con otras ciencias.



**Figura No.1. Relación entre la matemática y otras ciencias**

El presente trabajo se encuentra en la línea de investigación que intenta dilucidar acerca de la relación entre las prácticas sociales y la construcción de los conocimientos (Arrieta, 2003), una de las tesis centrales de esta línea sostiene que los conocimientos emergen de las prácticas de las comunidades, que viven ligados a dichas prácticas y, en este sentido, ligados a sus intencionalidades.

La comunidad de estudio es la conformada por los profesionales de la pesca, en la que se consideran tanto a los biólogos pesqueros como a los ingenieros pesqueros; siendo éstos el punto de partida. Al observar el currículo de la carrera del ingeniero pesquero, podemos darnos cuenta de que la modelación se estudia en diferentes momentos, sin embargo, es claro que al igual que en otras comunidades hay una separación de los conocimientos del aula con las prácticas de las comunidades como profesionistas y, por ende, de las intencionalidades, de esta manera ha nacido el mito del conocimiento por el conocimiento, el

conocimiento que vale por sí mismo. Ulloa y Arrieta, 2009

Esto no lleva a resaltar que, la escuela ha minimizado la creación matemática a partir de la experimentación en el laboratorio. El proceso de enseñanza y aprendizaje en asignaturas de primer año de la educación superior se ha transformado en un desafío para los docentes universitarios. La tecnología avanzada y el uso permanente de las redes sociales sugieren implementar en el aula, metodologías apoyadas en la Informática y la Comunicación, que le den más sentido al aprendizaje de las asignaturas de Ciencias Básicas. La química no escapa a esta problemática.

Por otro lado, entender el lenguaje de la química o aplicar sus leyes y principios, implica que los estudiantes manejen el vocabulario químico y los conceptos matemáticos que, de no tenerlos claros, será una limitante al momento de tratar de resolver un problema químico y, consecuentemente, en la aprobación de la asignatura. Lo anterior además se ve reflejado en el momento en que se desea utilizar la matemática para la solución de problemas de la química, como es el caso de la interpretación de resultados de análisis bromatológicos en especies marinas o de otro origen.

Por lo anterior uno de los objetivos del presente trabajo es analizar los resultados de composición de especies marinas a la luz de las herramientas matemáticas que se requieren, así como proponer alternativas para realizar interpretaciones y predicciones dejando como última opción el uso de software especializado ya que consideramos que en el campo laboral éste no está siempre presente.

La perspectiva teórica la socioepistemología es el marco en el cual se inscribe nuestra investigación. Concibe los fenómenos educativos como complejos, donde confluyen dimensiones, como lo son las epistemológica, la didáctica y la cognitiva en contextos sociales concretos. Esta perspectiva imprime en nuestras investigaciones lo situacional y lo multidimensional (Cantoral y Farfán, 2002; Cordero, 2002; Arrieta, 2003). La consideración de la dimensión social en nuestra perspectiva implica, modificar las demás dimensiones, en particular, lo social nos lleva a un plano donde se analiza los aprendizajes y la construcción de conocimiento por los actores en relación con el quehacer dentro de una sociedad. Desde esta perspectiva intentamos comprender y adecuar nuestro diseño a los diferentes contextos de acuerdo a sus necesidades. A partir de esta óptica, no pretendemos dar sólo una definición de lo que es lo multilínea, sino más bien tratar de caracterizar la naturaleza de esta red de prácticas y de herramientas.

#### **Antecedentes**

La investigación tiene diversos antecedentes, los trabajos acerca de la modelación como práctica social y las prácticas de análisis de los resultados de la composición de las especies son los principales. Uno de los aspectos fundamentales de esta línea de investigación consiste en situar el estudio de las prácticas de modelación en una comunidad, en un lugar y en un tiempo.

Las prácticas de modelación que se reportan en el estudio se enfocan en las que emergen en las actividades realizadas en las comunidades de biólogos e ingenieros pesqueros en interacción con fenómenos (físicos, químicos, sociales, etc.), conjeturando y realizando predicciones acerca de ellos utilizando modelos. En especial nos centraremos en los fenómenos en los que se trabaja con Modelos Exponenciales. Estas prácticas no solo se han ejercido históricamente, de la misma forma se ejercen en el plano profesional y en los

problemas cotidianos actuales.

Las prácticas de modelación que se han elegido se enfocan en las prácticas que se desarrollan en las comunidades de biólogos e ingenieros pesqueros en interacción con fenómenos (físicos, químicos, sociales, etc.), conjeturando y realizando predicciones acerca de ellos utilizando modelos.

Las actividades de modelación las distinguimos de quienes la usan con la finalidad de enseñar a modelar, a desarrollar teorías de modelación o hacer uso de ésta.

Reproducimos y analizamos prácticas de modelación con la intencionalidad explícita de desarrollar procesos de matematización en el aula. Nuestra perspectiva asume a las prácticas sociales como la base de nuestros diseños, en particular tomamos como base a las prácticas centradas bien en los modelos numéricos, bien en modelos gráficos o analíticos (Arrieta y Díaz, 2014).

En el aspecto de los análisis de composición (análisis bromatológicos) nos basaremos en los trabajos iniciados por Nieto, 2006, en las instalaciones de la Unidad Académica Escuela Nacional de Ingeniería Pesquera de la Universidad Autónoma de Nayarit, en los que hace una caracterización del contenido químico aproximado del músculo de 42 especies de peces que se comercializan en la región de San Blas, Nayarit. Es necesario resaltar que si bien Nieto inicia este tipo de trabajos en la Escuela Nacional de Ingeniería Pesquera, Munguía en 2004, presenta su tesis sobre un estudio de la composición química aproximada de *Scomberomorus sierra* (sierra) el cual publica en ese mismo año.

Nieto para visualizar la tendencia de agrupación de todas las especies de acuerdo con su composición química realizó dos clasificaciones

una mediante conglomerados considerando el vecino más próximo realizado con el software SPSS 10.0 y la de componentes principales mediante el programa XLSTAT, el análisis lo realiza con Sigma Stat 2.03, por su parte Munguía utiliza el programa estadístico Sigma Stat 2.03, primeramente se determinó la correlación existente entre los componentes químicos por el método de Pearson; luego de encontrar los componentes que tuvieron correlación se siguió a encontrar los modelos pertinentes en el mismo programa de los cuales se obtuvo la estadística descriptiva y los correspondientes gráficos.

### **Justificación**

Alimento se define, desde el punto de vista nutricional, como todo producto natural o transformado que por sus componentes químicos y características organolépticas puede ser ingerido para calmar el hambre, satisfacer el apetito y aportar los nutrientes que el organismo requiere para mantenerse sano, ya que, gracias a él, se desarrollan correctamente los procesos bioquímicos que sostienen la vida (López, 2012)

La comprensión de la importancia que tiene la composición química de los alimentos en la definición de su valor nutritivo data de fines del siglo pasado. Atwater, en 1894, estableció que "esta información era indispensable para mejorar el presupuesto familiar". Al mismo tiempo este conocimiento permitió el desarrollo de los primeros conceptos acerca de la relación entre la dieta y la salud de la población, iniciándose las primeras investigaciones sistemáticas tendientes a estudiar los requerimientos nutricionales del ser humano. El conocimiento de la composición química de los alimentos es el

primer paso esencial en el tratamiento dietético de la enfermedad y en cualquier estudio nutricional cuantitativo"

En la actualidad, una de las preocupaciones fundamentales de los expertos en nutrición es la de establecer relaciones entre el tipo de dieta consumida por la población y la prevalencia e incidencia de enfermedades crónicas no transmisibles. En este sentido, durante largo tiempo el énfasis radicó en conocer la composición de nutrientes de los alimentos y en el último tiempo cobra cada vez más importancia la composición química de los no nutrientes. El otro objetivo central sigue siendo la identificación de los nutrientes que están en déficit en la dieta de las poblaciones.

Las necesidades de información sobre composición de alimentos y las aplicaciones de las tablas en los distintos países, guardan una estrecha relación con las características de la situación alimentaria y nutricional de la población, con el desarrollo de la investigación en el tema y con la prioridad que asignan los gobiernos a la búsqueda de soluciones a los problemas nutricionales.

El uso de las tablas de composición química de los alimentos es muy amplio. A nivel nacional, permiten evaluar la adecuación de la disponibilidad nacional de alimentos con respecto a las necesidades nutricionales de la población, en términos de nutrientes, permitiendo además identificar eventuales deficiencias en dicha disponibilidad FAO 1992 .

En educación alimentaria y nutricional, las tablas son esenciales para expresar las recomendaciones nutricionales en guías

alimentarias que orienten a la población en la selección de una alimentación más saludable FAO 1996. Las tablas de composición de alimentos constituyen un material educativo por sí mismas. Es esencial que los estudiantes y profesionales de las carreras de la salud, educación y otras relacionadas con la formación de hábitos alimentarios saludables en la población, conozcan la composición de los alimentos y sean capaces de comprender su utilidad para cubrir las necesidades nutricionales.

El desarrollo de paquetes computacionales de fácil manejo para los usuarios que tengan acceso a este tipo de tecnología constituye un gran aporte a la velocidad y precisión de los análisis de la información recolectada, pero que ocurre en sitios en los que no se cuenta con este tipo de tecnología. Nuestra postura es proponer algún tipo de solución que esté al alcance para profesionistas con un bagaje básico de matemáticas.

Por otra parte, el poder relacionar mediante modelos matemáticos la relación entre los diferentes tipos de componentes puede ser de gran utilidad porque a partir de un análisis que además de ser relativamente rápido pero principalmente de bajo costo, pueden tenerse de manera aproximada los componentes en cuya determinación se requiere de equipo y reactivos costosos.

### **Metodología**

Una extensión útil en la regresión lineal es el caso en el que la variable dependiente ( $y$ ) es una función lineal de dos o más variables independientes ( $x_1, x_2, x_3, \dots$ ) de la forma:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n$$

Esta ecuación es útil particularmente cuando se ajustan datos experimentales como es el caso de la composición química de alimentos en donde la variable que se está analizando es función de otras dos o más variables.

En el caso bidimensional:

$$y = a_0 + a_1x_1 + a_2x_2$$

La representación de la regresión ya no es una línea recta ni una curva, sino un plano en el espacio, lo cual dificulta en cierto grado su representación, sin embargo, es posible utilizar el método de mínimos cuadrados para encontrar los coeficientes  $a_0$ ,  $a_1$  y  $a_2$  de con base en el procedimiento que se describe.

Se debe obtener la suma de los cuadrados de las diferencias o errores

$$Sr = \sum_{i=1}^n (y_i - a_0 - a_1x_1 - a_2x_2)^2$$

Derivando con respecto a cada uno de los coeficientes se tiene.

$$\frac{\partial Sr}{\partial a_0} - 2 \sum (y_i - a_0 - a_1x_1 - a_2x_2) = 0$$

$$\frac{\partial Sr}{\partial a_1} - 2 \sum x_1(y_i - a_0 - a_1x_1 - a_2x_2) = 0$$

$$\frac{\partial Sr}{\partial a_2} - 2 \sum x_2(y_i - a_0 - a_1x_1 - a_2x_2) = 0$$

Los coeficientes que generan la suma mínima de los cuadrados se obtienen al igualar a cero las derivadas parciales y se genera el sistema de ecuaciones:

$$\sum y_i = na_0 + \sum x_{1i}a_1 + \sum x_{2i}a_2$$

$$\sum x_{1i}y_i = \sum x_{1i}a_0 - \sum x_{1i}^2 a_1 + \sum x_{1i}x_{2i}a_2$$

$$\sum x_{2i}y_i = \sum x_{2i}a_0 + \sum x_{1i}x_{2i}a_1 + \sum x_{2i}^2 a_2$$

Las expresiones anteriores se pueden escribir en la forma matricial

$$\begin{bmatrix} n & \sum x_{1i} & \sum x_{2i} \\ \sum x_{1i} & \sum x_{1i}^2 & \sum x_{1i}x_{2i} \\ \sum x_{2i} & \sum x_{1i}x_{2i} & \sum x_{2i}^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_{1i}y_i \\ \sum x_{2i}y_i \end{bmatrix}$$

El coeficiente de correlación se calcula mediante la ecuación

$$r = \sqrt{\frac{St - Sr}{St}}$$

### Resultados

Utilizando los datos observados por Munguía (2004) para la sierra *Scomberomorus sierra*

MES	HUMEDAD g/100g	CENIZAS g/100g	PROTEINA g/100g
Enero	72.35	1.39	18.83
Febrero	72.36	1.29	19.17
Marzo	68.51	1.48	20.62
Abril	73.05	1.37	18.22
Mayo	72.06	1.48	18.69
Junio	74.04	1.47	15.66
Julio	73.80	1.30	17.90
Agosto	72.71	1.27	17.67
Septiembre	74.88	1.53	16.32
Octubre	72.80	1.59	16.90
Noviembre	67.93	1.20	19.99
Diciembre	69.38	1.51	18.55

El modelo tiene la forma de la ecuación

$$y = a_0 + a_1x_1 + a_2x_2$$

El sistema que se debe plantear es

$$\begin{bmatrix} n & \sum x_{1i} & \sum x_{2i} \\ \sum x_{1i} & \sum x_{1i}^2 & \sum x_{1i}x_{2i} \\ \sum x_{2i} & \sum x_{1i}x_{2i} & \sum x_{2i}^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_{1i}y_i \\ \sum x_{2i}y_i \end{bmatrix}$$

En consecuencia, se debe construir una tabla 2.

Lo que se representa como

$$\begin{bmatrix} 12 & 863.87 & 15.58 \\ 863.87 & 62243.09 & 1215.7517 \\ 15.58 & 1215.7517 & 23.9068 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 200.62 \\ 15702.4937 \\ 306.6134 \end{bmatrix}$$

Resolviendo se tiene:

$$a_0 = 1.73309$$

$$a_1 = -0.033526$$

$$a_2 = 13.4009$$

El modelo es:  $y = 1.73309 - 0.033526x_1 + 13.4009x_2$

Es decir:

$$Proteína = 1.73309 - 0.033526 * Humedad + 13.4009 * Ceniza$$

Tabla No. 2

yi	x1i	x2i	x1i * x2i	x1i^2	x2i^2	x1i*yi	x2i*yi
18.83	72.35	1.39	100.5665	5234.523	1.9321	1362.3505	26.1737
19.17	72.36	1.29	93.3444	5235.97	1.6641	1387.1412	24.7293
20.62	68.51	1.48	101.3948	4693.62	2.1904	1412.6762	30.5176
18.22	73.05	1.37	100.0785	5336.303	1.8769	1330.971	24.9614
18.69	72.06	1.48	106.6488	5192.644	2.1904	1346.8014	27.6612
15.66	74.04	1.47	108.8388	5481.922	2.1609	1159.4664	23.0202
17.90	73.8	1.30	95.94	5446.44	1.69	1321.02	23.27
17.67	72.71	1.27	92.3417	5286.744	1.6129	1284.7857	22.4409
16.32	74.88	1.53	114.5664	5607.014	2.3409	1222.0416	24.9696
16.9	72.8	1.59	115.752	5299.84	2.5281	1230.32	26.871
19.99	67.93	1.2	81.516	4614.485	1.44	1357.9207	23.988
18.55	69.38	1.51	104.7638	4813.584	2.2801	1286.999	28.0105
200.62	863.87	15.58	1215.7517	62243.09	23.9068	15702.4937	306.6134

El cálculo del coeficiente de correlación nos da:

Para el cálculo del coeficiente de correlación, se tiene:

$(y_i - a_0 - a_1x_{1i} - a_2x_{2i})^2$	$(y_i - y_{prom})^2$
0.801499599	0.35024876
6.634180831	0.86828512
1.823699699	5.67305785
0.332642062	0.00033058
0.212095655	0.20413967
10.8250736	6.64702149
1.488299474	0.11436694
1.837224262	0.32283058
11.6011093	3.67942149
13.68872871	1.79073058
19.83144607	3.06886694
1.193370794	0.09723058
<b>Suma</b>	<b>70.26937006</b>
	<b>22.8165306</b>

$$r = \sqrt{\frac{St - Sr}{St}} = \sqrt{\frac{70.26937006 - 22.8165306}{70.26937006}} = \sqrt{0.67529}$$

$r = 0.821$  el cual indica que es un valor medianamente aceptable por su cercanía a 1.0.

Utilizando la calculadora TI - Nspire CX, el modelo que se obtiene es:

$$y = 58.5519 - 0.4952x_1 - 3.32985x_2$$

Es decir:

$$\text{Proteína} = 58.5519 - 0.4952 \cdot \text{Humedad} - 3.32985 \cdot \text{Ceniza}$$

Con un coeficiente de correlación  $r = 0.843647$

### **Discusión**

El modelo obtenido con base en el software difiere del obtenido mediante la utilización de ecuaciones no obstante los coeficientes de correlación son muy semejantes, siendo de 0.843647 con el uso de software y de 0.821 en el caso del cálculo analítico.

El modelo de regresión múltiple de predicción de proteína muestra un comportamiento cercano a los resultados de los análisis, por lo que su utilización debe ser con fines de tener un aproximado, por lo que si se considera un factor de corrección el modelo se ajusta más a los valores reales.

La valoración y uso de los modelos es responsabilidad de quien los utilice. Cabe mencionar que con su utilización existe el ahorro de reactivos y tiempo en la realización de análisis, por otro lado, hay que considerar las variables consecuencia a la localidad donde fueron tomados los datos, la biodiversidad que permite la zona no es necesariamente la misma en todas partes, de ahí que esto sea una propuesta para cual se deben de tomar sus propios datos y desde ahí realizar los análisis, que permitan tomar las decisiones referentes a éste fenómeno en particular.

### **Conclusiones**

En primera instancia es posible observar la relación existente entre la humedad y la proteína, como se puede observar en el trabajo de Munguía existe una relación derivada de la alimentación del organismo en cuestión, así como la temporada del año en

que fue observada, esto nos permite que al reconocer el impacto de una variable específica en el organismo se pueda buscar una forma de sustituir la alimentación para que se obtenga lo que más convenga al investigador.

Por otro lado, este estudio nos permite analizar de manera crítica, situaciones propias de una comunidad (la de ingenieros pesqueros), en la que, la intervención de la matemática a través de modelos permite un desempeño óptimo, y al mismo tiempo reconocer el impacto que tiene el clima en el desarrollo de una especie en específico.

Para este fin, un análisis de múltiples variables permite, en un primer estadio reconocer los valores de aproximación en un modelo general, pero a su vez con éstos datos modificar las condiciones del proceso para recoger resultantes distintos y llegar a un consumo óptimo de recursos.

Para cerrar se puede reconocer, que el uso de herramientas tecnológicas mejora de sobremanera la interacción entre el fenómeno y los miembros de la comunidad, ya que, es a través de ésta que se puede localizar que variable en la cual se puede intervenir para la manipulación de un fenómeno con las ventajas que esto conlleva.

### **Referencias Bibliográficas**

Arrieta, J.; Díaz, L. (2014). Una perspectiva de la modelación desde la Socioepistemología. *Revista Latinoamericana de Investigación en Matemática Educativa* (2015) 18 (1): 19-48

Atwater, W. (1894). Foods: Nutritive Value and Cost. Secretary of Agriculture. Washington.

Cordero, F. (2002). Lo social en el conocimiento matemático: los argumentos y la reconstrucción de significados. En J. Delgado (Ed.), Acta Latinoamericana de Matemática Educativa (Vol. 16, Tomo I, pp. 45-50). Chile.

Cantoral, R. y Farfán, R. (2002). Sur la sensibilité a la contradiction en mathématiques; l'origine de l'analyse complexe. Recherches en Didactique des mathématiques. Vol. 22, Núm. 2.

FAO. (1992). AGROSTAT. Hojas de balance de alimentos

FAO/WHO. (1996). Preparation and use of food-based dietary guidelines. Report of a joint

FAO/WHO consultation. Nicosia, Cyprus. Nutrition Programme. WHO Geneva 1996

López, V. (2012). Composición química de los alimentos. Red Tercer Milenio. México

Munguía, J. (2004). Análisis químico proximal de *Scomberomorus sierra* durante el periodo de Enero a Diciembre de 2003 en San Blas Nayarit. Tesis de Licenciatura no publicada. Universidad Autónoma de Nayarit

Nieto, J. (2006). Análisis proximal de peces comerciales de la región de San Blas Nayarit. Tesis de Maestría no publicada. Universidad Autónoma de Nayarit - Universidad de Guadalajara, México

Ulloa, J.; Arrieta, J. (2009). Los modelos exponenciales: construcción y deconstrucción. En P. Lestón (Ed), Acta Latinoamericana de Matemática Educativa 22, 479-488. México: Comité Latinoamericano de Matemática Educativa.

Ulloa, J.; Arrieta, J.; Benítez, A. 2015. Alternativas para la elaboración de modelos matemáticos. Acta Pesquera 1 impresa, pp 42, 57

