

Regresión multilineal y química de alimentos

Multilinear Regression and Food Chemistry

José Trinidad Ulloa Ibarra¹, José Trinidad Nieto Navarro¹,
María Inés Ortega Arcga¹, Juan Felipe Flores Robles¹, Nidia D. Uribe Olivares

¹ Universidad Autónoma de Nayarit
² CBETIS 100

Recibido: 30 de septiembre de 2019
Aceptado: 12 de noviembre de 2019

Resumen.

Continuando con el estudio de los trabajos de modelación que realizan los miembros de la comunidad de profesionales de la pesca, se aborda en este caso, la modelación multilineal para situaciones de la química de alimentos y cuyo fin es mostrar que las herramientas matemáticas nos permiten establecer modelos matemáticos que relacionan diferentes componentes químicos con una aproximación aceptable y que nos permite inferir la cantidad de componentes con base en las determinaciones proximales de bajo costo y tiempo, lo que redundará en un ahorro económico y de tiempo que en el caso de los productos perecederos es algo de gran importancia. Seguimos tomando como sustento teórico a la socioepistemología que analiza 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, grasa

Abstract

Continuing with the study of modeling work carried out by members of the community of fisheries professionals, in this case, multilinear modeling for situations of food chemistry is addressed and whose purpose is to show that mathematical

tools allow us to establish mathematical models that relate different chemical components with an acceptable approximation and that allow us to infer the quantity of components based on the proximal determinations of low cost and time, which results in an economic and time saving than in the case of perishable products. It is something of great importance. We continue to take the socio-epistemology that analyzes the educational phenomenon as a theoretical support considering the environment in which it is carried out, so we maintain that we carry out a study in a community whose results when being transferred to the classroom benefit the two poles in which it interacts.

Key words: Multilinear regression, proximal tests, food, fat.

Introducción.

Las ciencias son un conjunto de conocimientos adquiridos por la humanidad, una necesidad del ser humano para su progreso y desarrollo, son un acto creativo del individuo. La gran mayoría de estas ciencias están relacionadas con la ciencia lenguaje del universo: la matemática. Ésta les ha aportado criticidad y les ha permitido el desarrollo de grandes teorías y aplicaciones; basta estudiar alguna de ellas en particular para ver su huella plasmada en el fantástico concierto de sus teorías, que da muestra del profundo poder de creación que tiene la figura más compleja del universo: el hombre (Rodríguez, 2011).

En todas las ciencias está presente la matemática y por tanto puede y debería usarse la relación matemática-ciencias como recurso didáctico en cualquier nivel educativo, es en este sentido en que cobra relevancia la propuesta socioepistemológica ya que se debe ofrecer al estudiante un acercamiento a otras ciencias desde la matemática y viceversa, percibiendo que todos los campos del saber están relacionados de alguna manera; mostrar la profunda transdisciplinariedad de las ciencias, enseñar matemática como si estuviesen aisladas es una distorsión del conocimiento.

Convendría enseñar Matemática yendo más allá de las propias Matemática: considerando sus relaciones y buscando su sintonía con las corrientes principales del pensamiento. Esta nueva actitud motivaría a los estudiantes, crearía nuevas aplicaciones y abriría nuevas vías de debate (Gómez, 2002).

El conocimiento no está aislado, sino que se encuentra intrínsecamente relacionado en sus diferentes componentes por lo que sus partes deben ser contextualizadas en el marco de determinado problema concreto que desafíe al sujeto y que le permita retomar un aprendizaje significativo. Se trata de un aprendizaje que, para el logro de su objetivo en cuanto a resolución de un problema, requiere en su aplicación del tránsito desde el problema de realidad que se pretende resolver, al reconocimiento y fortalecimiento de las categorías lógicas-matemáticas que involucra dicha resolución.

A diferencia de lo que ocurre en el contexto escolar, en los contextos laborales -o de la vida cotidiana- se presentan situaciones problemáticas menos estructuradas y más difusas respecto de las variables que deben seleccionarse para un correcto planteo y eficaz resolución. Estos últimos contextos requieren por parte de los adultos -sus protagonistas- el desarrollo o fortalecimiento de habilidades que permitan:

- Buscar, analizar y seleccionar datos disponibles o inferidos.
- Organizar los datos como información.
- Formular hipótesis que permitan traducir al lenguaje matemático el problema presentado.
- Diseñar variables que contribuyan a explicar el fenómeno o el problema presentado.
- Establecer razonamientos y relaciones que hagan posible plantear o diagnosticar el problema.
- Establecer relaciones matemáticas que permitan orientar la decisión sobre la mejor forma de resolver el problema.
- Verificar sobre la situación problemática real si la solución matemática es aceptable.

Las personas interactúan con el mundo coti-

diano mediante el uso de lenguajes que permiten el desarrollo de determinadas capacidades. En particular, el lenguaje matemático, a diferencia de otros, posibilita el desarrollo y fortalecimiento de las siguientes capacidades:

- Pensar y razonar. Incluye plantear formas de identificar, discriminar, diferenciar, cuantificar, buscar, entender y manipular el rango y los límites de ciertos conceptos matemáticos.
- Argumentar. Incluye establecer y/o evaluar cadenas de argumentos lógico-matemáticos de diferentes tipos; desarrollar procedimientos intuitivos, y construir y expresar argumentos matemáticos.
- Comunicar. Involucra la habilidad de expresarse, tanto en forma oral como escrita, sobre asuntos con contenido matemático. Implica también entender las aseveraciones orales y escritas expresadas por otros sobre los mismos temas.
- **Modelar.** Traduce la "realidad" -o la situación problemática identificada- a un modelo matemático, el cual deberá ser validado a través del análisis y la crítica de este y de sus resultados, estableciendo un monitoreo y control del proceso de modelado. El modelo y sus resultados deberán ser comunicables y permitir el señalamiento de sus limitaciones y restricciones.
- Plantear y resolver problemas. Comprende las habilidades de formular y definir diferentes clases de problemas matemáticos, y de resolverlos mediante el uso de diversos métodos, estrategias y algoritmos.
- Representar. Incluye la habilidad de codificar y decodificar, traducir, interpretar y distinguir entre diferentes tipos de representaciones de objetos y situaciones matemáticas. Esta habilidad contempla la elección entre las diferentes formas de representación y sus interrelaciones de acuerdo con la situación y el propósito particular.
- Utilizar lenguaje y operaciones simbólicas, formales y técnicas. Comprende la habilidad de decodificar e interpretar lenguaje formal y simbólico, y entender su relación con el lenguaje coloquial; traducir desde el lenguaje

al lenguaje simbólico/formal; manipular proposiciones y expresiones que contengan símbolos y fórmulas; realizar cálculos, utilizar variables y resolver ecuaciones.

- Utilizar ayudas y herramientas. Involucra la habilidad de conocer y ser capaz de utilizar diversas ayudas y herramientas, incluidas las tecnologías de la información y las comunicaciones.

Química y Matemáticas

La relación química - matemáticas es tan antigua como la humanidad, los alquimistas usaron las matemáticas solo con propósitos mágicos; parece que nunca desarrollaron modelos matemáticos para explicar fenómenos químicos. Las únicas herramientas matemáticas usadas por los alquimistas fueron de tipo aritmético y geométrico, únicos campos de las matemáticas bien desarrollados en aquellos tiempos.

El primer intento de "matematizar" la química se debe a Alexander Crum Brown (1838 -1922) un químico orgánico escocés, subestimado en la historia de la química. En un artículo (¡de 19 líneas!), representaba: compuestos químicos 'operandos' y los procesos químicos 'operadores'. Arthur Cayley (1821-1895) desarrolló las matrices que más tarde han resultado esenciales para el progreso de la química cuántica y la química matemática. Heisenberg redescubrió las matrices cuando desarrollo la mecánica de matrices. De los 342 artículos publicados por James Sylvester (1814-1897) solo dos están dedicados a la química (1878) y son fundamentales en química matemática ('química algebraica'). En un artículo en Nature introdujo el término chemigraph (graph = grafo) para la notación gráfica química.

La matematización de la Química servirá para:

- Establecer las bases teóricas (fundamentales).
- Interpretar más fácilmente los resultados.
- Aumentar el poder de predicción.

La Química matemática es el área científica que se encarga de las aplicaciones de las matemáticas en la química. Se trata de usar instrumentos matemáticos que ayuden a Química Matemática: Se trata de usar instrumentos matemáticos que ayuden a modelizar los procesos químicos y no se debe confundir con la química computacional

Uno de los objetivos del presenta 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.

Antecedentes

La investigación tiene diversos antecedentes, los principales antecedentes son los trabajos acerca de la modelación como práctica social y las prácticas de análisis de los resultados de la composición de las especies, el primero desarrollado en diferentes centros de investigación y docencia y los segundos referidos a los trabajos realizados en la Unidad Académica Escuela Nacional de Ingeniería Pesquera. 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.

Sobre la modelación aplicada a relacionar los principales componentes de las especies marina, resaltamos los trabajos realizados por Nieto, 2006 y los trabajos posteriores de tesis en la licenciatura, tales como el desarrollado por Munguía en 2004, Hermosillo y Caamal, 2005; se tiene registros de los trabajos de Ramos, 2009 y de Mónico, 2010. En todos ellos, el análisis matemático para relacionar los componentes se basó en la utilización de software tal como SPSS, Sigma Stat y Excel con lo que se determinaron los gráficos y la correlación entre dos o más componentes químicos

Justificación

La principal técnica de determinación del tiempo de vida útil que se ha utilizado es el estudio de vida de anaquel, la cual es costosa, puesto que se deben realizar pruebas de calidad en un periodo de tiempo prolongado. Actualmente la determinación se realiza por medio de pruebas aceleradas que permiten, mediante modelación matemática establecer con exactitud la fecha de vencimiento del alimento evaluado.

Recientemente la empresa Kraft se ha centrado en el desarrollo de una tecnología de análisis/rastreo que utiliza las matemáticas para identificar nuevos compuestos con efectos específicos sobre la salud. Kraft, en colaboración con Medisyn Technologies, aclara que con este estudio esperan poder acortar el tiempo y los costes derivados de la investigación y el desarrollo de los ingredientes funcionales. La empresa ha utilizado esta tecnología para escanear y registrar cientos de miles de compuestos para encontrar los modelos que se adapten a cada tipo de actividad biológica deseada.

Este método es, en su esencia, el opuesto a la forma tradicional de encontrar este tipo de ingredientes activos. En lugar de desarrollar propiedades a través de compuestos ya conocidos, se establece primeramente un modelo matemático que responde a una propiedad o actividad específica concreta, y luego se aplica este modelo a las bases de datos creadas con los miles de compuestos para tratar de encontrar uno que se acople a dicho modelo. De esta forma se consigue relacionar una propiedad previamente establecida (definida por un modelo matemático), con un compuesto específico, almacenado en la base de datos.

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 res-

pecto a las necesidades nutricionales de la población, en términos de nutrientes, permitiendo además identificar eventuales deficiencias en dicha disponibilidad FAO 1992.

Estamos en cierto sentido en el terreno de la Química matemática que es el área científica que se encarga de las aplicaciones de las matemáticas en la química. Se trata de usar instrumentos matemáticos que ayuden a modelar los procesos químicos y no se debe confundir con la química computacional. La matematización de la Química servirá para: Establecer las bases teóricas (fundamentales); Interpretar más fácilmente los resultados; Aumentar el poder de predicción. Esto último por medio de la modelación

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 me-

diante 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	CENI- ZAS g/100g	E. ETÉ- REO g/100g
Enero	72.35	1.39	7.43
Febrero	72.36	1.29	7.18
Marzo	68.51	1.48	9.39
Abril	73.05	1.37	7.36
Mayo	72.06	1.48	7.77
Junio	74.04	1.47	8.83
Julio	73.80	1.30	7.00
Agosto	72.71	1.27	8.35
Septie	74.88	1.53	7.27
Octubre	72.80	1.59	8.71
Noviem	67.93	1.20	10.88
Diciemb	69.38	1.51	10.56

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 como la siguiente:

	yi	x1i	x2i	x1i * x2i	xi1^2	x2i^2	x1i*yi	x2i*yi
	7.43	72.35	1.39	100.5665	5234.5225	1.9321	537.5605	10.3277
	7.18	72.36	1.29	93.3444	5235.9696	1.6641	519.5448	9.2622
	9.39	68.51	1.48	101.3948	4693.6201	2.1904	643.3089	13.8972
	7.36	73.05	1.37	100.0785	5336.3025	1.8769	537.648	10.0832
	7.77	72.06	1.48	106.6488	5192.6436	2.1904	559.9062	11.4996
	8.83	74.04	1.47	108.8388	5481.9216	2.1609	653.7732	12.9801
	7	73.8	1.3	95.94	5446.44	1.69	516.6	9.1
	8.35	72.71	1.27	92.3417	5286.7441	1.6129	607.1285	10.6045
	7.27	74.88	1.53	114.5664	5607.0144	2.3409	544.3776	11.1231
	8.71	72.8	1.59	115.752	5299.84	2.5281	634.088	13.8489
	10.88	67.93	1.2	81.516	4614.4849	1.44	739.0784	13.056
	10.56	69.38	1.51	104.7638	4813.5844	2.2801	732.6528	15.9456
Σ =	100.73	863.87	16.88	1215.7517	62243.0877	23.9068	7225.6669	141.7281
Prom =	8.3941666 = 67							

Lo que se representa como

$$\begin{bmatrix} 12 & 863.87 & 16.88 \\ 863.87 & 62243.09 & 1215.7517 \\ 16.88 & 1215.7517 & 23.9068 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 100.73 \\ 7225.6669 \\ 141.7281 \end{bmatrix}$$

$$r = \sqrt{\frac{19.22389167 - 6.35200144}{19.22389167}} = 0.82119851$$

Coefficiente de correlación múltiple = 0.82198151

Resolviendo se tiene:

$$a_0 = 41.6483074$$

$$a_1 = -0.5007406$$

$$a_2 = 1.98608596$$

El modelo es: $y = 41.6483074 - 0.5007406x_1 + 1.98608596x_2$

Es decir:

$$Grasa = 41.6483074 - 0.5007406 * Humedad + 1.98608596 * Ceniza$$

El cálculo del coeficiente de correlación nos da:

$r = 0.82198151$ 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 se muestra en la siguiente figura.

$$y = 41.6483 - 0.5007406x_1 - 1.98609x_2$$

Es decir:

$$Grasa = 41.6483 - 0.5007406 * Humedad + 1.986096 * Ceniza$$

Con un coeficiente de correlación $r = 0.821982$

The process is shown in six sequential screenshots:

- Screenshot 1:** A spreadsheet with data in columns A, B, and C. Row 1: 7.43, 72.35, 1.39. Row 2: 7.18, 72.36, 1.29. Row 3: 9.39, 68.51, 1.48. Row 4: 7.36, 73.05, 1.37. Row 5: 7.77, 72.06, 1.48.
- Screenshot 2:** The 'Estadística' menu is open, showing options like 'Cálculos estadísticos', 'Distribuciones', 'Intervalos de confianza', and 'Pruebas estadísticas'.
- Screenshot 3:** The 'Regresión lineal múltiple...' option is selected in the 'Regresión' submenu.
- Screenshot 4:** The 'Regresión lineal múltiple' dialog box is shown with 'Núm. de vars ind:' set to 2. The spreadsheet background shows the data with labels 'gg', 'hh', and 'cc' above columns A, B, and C respectively.
- Screenshot 5:** The 'Regresión lineal múltiple' dialog box is shown with 'Lista Y:' set to 'a[]', 'Lista X1:' set to 'b[]', 'Lista X2:' set to 'c[]', 'Guardar EcnReg en:' set to 'f5', and '1ª columna de resultado:' set to 'd[]'.
- Screenshot 6:** The spreadsheet shows the results of the regression. The formula bar shows '=MultReg(' and the spreadsheet contains the following data:

	hh	cc	Título	Regresió...
1:	72.35	1.39		
2:	72.36	1.29	RegEqn	$b_0 + b_1 * x_1 + b_2 * x_2$
3:	68.51	1.48	b0	41.6483
4:	73.05	1.37	b1	-0.500741
5:	72.06	1.48	b2	1.98609

.Discusión

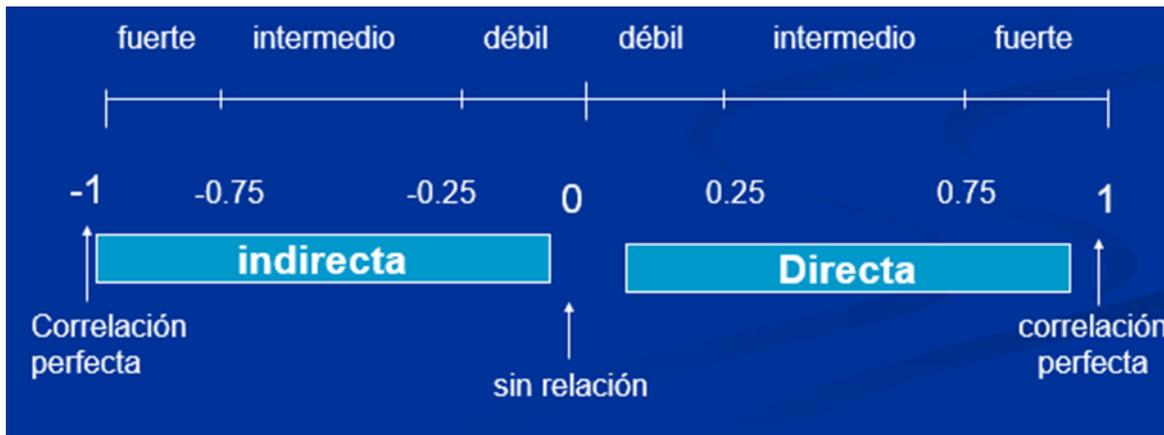
Puede establecerse que el modelo obtenido por cualquiera de los dos procedimientos es igual por lo que el método de obtención puede ser el que sea mejor comprendido por quien lo está proponiendo y esto dependerá de las herramientas de que disponga, ya que en nuestro caso la utilización de la calculadora es un procedimiento común, para este y otro tipo de modelos.

El modelo de regresión múltiple de predicción de grasa (extracto etéreo) 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.

No obstante, y de tomando como base la definición de Coeficiente de Correlación:

El valor de r denota la fuerza de la asociación como se ilustra en el siguiente diagrama.



La valoración y uso de los modelos es responsabilidad de quien los utilice (Ulloa, Nieto, Ortega, Flores y Arrieta, 2019).

Conclusiones

Con base en los resultados obtenidos se encuentra que es posible realizar predicciones acerca del contenido de grasa basándose para ello en la determinación de la humedad y la ceniza, pero es necesario considerar lo expuesto por Munguía, 2004, en el sentido de que la composición química de los peces está influenciada por la alimentación, la temporada del año, y otros factores del hábitat propio de la especie en cuestión.

Este ejercicio del establecimiento de modelos multilineales que es propio de la comunidad en estudio en donde la matemática y el uso de herramientas tecnológicas, tales como calculadora o software son de gran importancia para poder llegar a una buena aproximación y por lo tanto contar con una base para realizar predicciones que permitan ahorro de dinero y de tiempo.

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 estos datos modificar las condiciones del proceso para recoger resultantes distintos y llegar a un consumo óptimo de recursos (Ulloa, et al. 2019).

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
- FAO. (1992). AGROSTAT. Hojas de balance de alimentos
- Gómez, J. (2002). De la enseñanza al aprendizaje de las matemática. Barcelona: Paidós.
- Hermosillo, M.; Caamal, J. (2005). Composición química proximal en la almadraba durante el periodo febrero 2002 - 2003 en la Bahía de Matanchén. Tesis de Licenciatura no publicada. Universidad Autónoma de Nayarit.
- Mónico, L. (2010). Modelos matemáticos asociados a los análisis proximales del *Centropomus robalito* (Constantino) en San Blas, Nayarit. Tesis de Licenciatura no publicada. Universidad Autónoma de Nayarit.
- 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
- Ramos, D. (2009). Modelos matemáticos asociados a la composición proximal de *Orthopristis Chalceus* (Burrito) y *Cynoscion Parvipinnis* (Corvina). Tesis de Licenciatura no publicada. Universidad Autónoma de Nayarit.
- Rodríguez, M. (2011). La matemática y su relación con las ciencias como recurso pedagógico. *Números revista didáctica de las Matemáticas*. Vol. 77, pp. 35 - 49.
- Ulloa, J.; Nieto, J.; Ortega, M.; Robles, F.; Arrieta, J. (2019). Regresión multilineal como apoyo a los análisis proximales. *Acta Pesquera Volumen 5, No. 9*. Universidad Autónoma de Nayarit

