DISEÑO DE SIMULADORES Y APLICATIVOS DE MODELOS DE REGRESIÓN LINEAL Y NO LINEAL

CATALINA DEL PILAR MURCIA FLOREZ

FREDY PEÑA ACUÑA

UNIVERSIDAD PEDAGÓGICA NACIONAL
FACULTAD DE CIENCIA Y TECNOLOGÍA
DEPARTAMENTO DE MATEMÁTICAS
ESPECIALIZACIÓN EN EDUCACIÓN MATEMÁTICA
BOGOTÁ
2012

DISEÑO DE SIMULADORES Y APLICATIVOS DE MODELOS DE REGRESIÓN LINEAL Y NO LINEAL

CATALINA DEL PILAR MURCIA FLOREZ

2012182026

FREDY PEÑA ACUÑA

20121820

Trabajo de grado para optar al título de Especialista en Educación Matemática

Asesor:

BENJAMÍN SARMIENTO

UNIVERSIDAD PEDAGÓGICA NACIONAL
FACULTAD DE CIENCIA Y TECNOLOGÍA
DEPARTAMENTO DE MATEMÁTICAS
ESPECIALIZACIÓN EN EDUCACIÓN MATEMÁTICA
BOGOTÁ
2012

NOTA DE ACEPTACION

FIRMA DEL JURADO

FIRMA DEL JURADO

Bogotá, Noviembre de 2012

Para todos los efectos, declaramos que el presente trabajo es original y de nuestra total autoría; en aquellos casos en los cuales hemos requerido del trabajo de otros autores o investigadores, hemos dado los respectivos créditos.

A nuestros padres, hermanos y demás familia,

por su constante apoyo y motivación

AGRADECIMIENTOS

Debemos agradecer de manera especial y sincera al Profesor Benjamín Sarmiento por aceptarnos, para realizar esta propuesta bajo su orientación. Su ayuda y confianza en nuestro trabajo y su aptitud para orientar nuestras ideas fueron una contribución invaluable en el desarrollo de este proyecto. Las ideas propias, siempre enmarcadas en su rigor y profesionalismo, han sido la clave para del buen trabajo que hemos realizado juntos, el cual no se puede concebir sin su constante y oportuna colaboración.

También queremos agradecer muy especialmente, a nuestros Padres y Hermanos, y en general a cada miembro de nuestras familias, por su colaboración y valiosos aportes en la realización del proyecto.

A todos ellos, y a todas las personas que no nombramos aquí y que de una u otra manera aportaron para que este trabajo llegara a buen término, infinitas gracias.

1. Información General	
Tipo de documento	Tesis de grado
Acceso al documento	Universidad Pedagógica Nacional. Biblioteca Central
Titulo del documento	DISEÑO DE SIMULADORES Y APLICATIVOS DE MODELOS DE REGRESIÓN LINEAL Y NO LINEAL
Autor(es)	MURCIA FLOREZ, Catalina del Pilar. PEÑA ACUÑA, Fredy
Director	SARMIENTO, Benjamín
Publicación	Bogotá, 2012, 66 p.
Unidad Patrocinante	Universidad Pedagógica Nacional
Palabras Claves	Educación, Didáctica, Educación estadística, didáctica de la estadística, modelos de regresión, regresión lineal, regresión no lineal, material educativo multimedia.

2. Descripción

En la presente propuesta se describe y caracteriza el diseño de simuladores y aplicativos de los modelos de regresión lineal y no lineal, para ilustrar y comparar, a partir de problemas, los diferentes modelos de regresión lineal y no lineal abordados en grado 11. De esta manera se pretende que dicho material interactivo apoye los procesos de enseñanza y aprendizaje del objeto estadístico nombrado.

3. Fuentes

Se hacen referencia a 19 fuentes bibliográficas en las que se destacan:

Batanero, C. (2001). Didáctica de la Estadística. Granada: Grupo de Investigación en Educación

Estadística. ISBN 84-699-4295-6.

Batanero, C. &. (2000). El Papel de los Proyectos en la Enseñanza y Aprendizaje de la Estadística. Aspectos didácticos de las matemáticas , 125-164.

Batanero, C. (s.f.). ¿Hacia dónde va la educación estadística? Granada: Universidad de Granada.

Batanero, Estepa & Godino. (1991). Análisis exploratorio de datos: sus posibilidades en la enseñanza secundaria. Suma , 25-31.

Fernández, F., Soler, N. y Sarmiento, B. (2008). Conocimiento estadístico y probabilístico de profesores de educación básica y media (Reporte de investigación). Bogotá: Universidad Pedagógica Nacional.

Godino, J. (1995). ¿Qué aportan los ordenadores a la enseñanza de la estadística? UNO , 45 - 56.

López y otros. (s.f.). Excel como una herramienta asequible en la enseñanza de la Estadística. Recuperado el 20 de Septiembre de 2012, de http://campus.usal.es/~teoriaeducacion/rev_numero_07/n7_art_lopez_lagunes_herrera.htm

Ministerio de Educación Nacional. (2006). Estándares Básicos de Competencias en Matemáticas. Bogotá.

Ministerio de Educación Nacional. (1998). Lineamiento Curriculares de Matemáticas. Bogotá: MEN.

Sanchez, F. (1999). Significado de la correlación y regresión para los estudiantes universitarios. Tesis . Granada, España: Universidad de Granada.

4. Contenidos

El documento se divide en 4 capítulos nombrados:

CAPÍTULO 1. PROCESO DE INDAGACIÓN. Muestra el proceso de construcción y planificación realizado para el desarrollo del trabajo, que inicia dando una mirada a la reconstrucción teórica de los modelos de regresión lineal y no lineal, realizada a partir del estudio de investigaciones, tesis de diferentes niveles de educación profesional, artículos de revistas, entre otros; los cuales son presentados como los antecedentes de la propuesta.

Posteriormente, se encuentra la justificación del trabajo y se delimita y describe el tema de interés, del cual surgen las preguntas específicas de interés para el estudio, las cuales constituyen una pregunta global que guía el desarrollo de la investigación, y por medio de estas se plantean los objetivos (general y específicos) del trabajo de grado. Seguidamente, se da a conocer un plan de acción que busca abordar el tema de interés planteado, donde se presenta la población a la que está dirigido el trabajo y especificando las fases de la indagación durante el tiempo de su realización.

CAPÍTULO 2. REFERENTES TEÓRICOS Y DIDÁCTICOS. Da a conocer los aspectos que intervienen en la construcción y caracterización del marco teórico y didáctico, el cual compone los desarrollos a través del tiempo de investigaciones a partir de tres dimensiones relacionadas con lo epistémico, didáctico y curricular, referente a los modelos de regresión lineal y no lineal

CAPÍTULO 3. METODOLOGÍA Y DISEÑO DE APLICATIVOS Y SIMULADORES. Muestra un recuento y explicación de los elementos que se tienen en cuenta durante el desarrollo y diseño de la propuesta, realizando una descripción de los simuladores y aplicativos construidos, a través de la identificación de las partes y componentes del software, para después relacionar dichos aplicativos con los estándares en matemáticas de grado 11º referidos al objeto de estudio. Por último, en este capítulo, se dan a conocer algunos ejemplo de las situaciones problemas que se pueden abordar haciendo uso de los aplicativos y simuladores.

CAPÍTULO 4. REFLEXIONES Y CONCLUSIONES. Da a conocer las reflexiones en torno a les tres dimensiones trabajadas en los antecedentes y marco teórico. Además se realiza una reflexión respecto a los aportes que el trabajo brinda en general a los profesores de matemáticas y en particular a los autores de la presente propuesta.

5. Metodología

La metodología utilizada en la presente propuesta no está sustentada bajo un referente teórico determinado, ya que corresponde un proceso de indagación y no a uno de investigación. Bajo estas condiciones, se establece unos elementos metodológicos, que están ligados a la forma como se construye el trabajo y los tiempos en los cuales se ejecuta cada uno de ellos.

6. Conclusiones

- Las simulaciones y aplicativos son propuestos como una herramienta de ayuda para los estudiantes, la cual facilita el análisis de la situación planteada a través de distintos tipos de representación. El primero de ellos corresponde a la ecuación de la función involucrada, permitiendo que el estudiante reconozca la representación analítica de la regresión de cada situación; el segundo es la representación gráfica de la función, que no sólo se presenta para relacionar la representación analítica con ésta, sino para que el estudiante determine el modelo que mejor se ajusta al conjunto de datos y por tanto a la situación.

- El software evita que los estudiantes realicen y desarrollen algoritmos extensos y procesos repetitivos, con los cuales posiblemente no logren construir el concepto, ni le den sentido a la situación y mucho menos a la solución de la misma.

- El diseño y construcción de simulaciones y aplicativos acompañados de actividades que pretende ofrecer a la comunidad de educadores matemáticos, y en particular estadísticos, herramientas y materiales educativos que apoyen el desarrollo y ejecución de los planes de estudio propuestos en cada institución educativa y de los estándares matemáticos y políticas nacionales, en la educación media.

- Respecto a la ejecución de los planes de estudio, los aplicativos permiten orientar los procesos de enseñanza – aprendizaje, los cuales facilitan el trabajo del profesor y del estudiante y se constituyen como una alternativa para abordar temáticas y competencias que involucren conceptos abstractos, los cuales son de difícil entendimiento sin ayudas didácticas y gráficas oportunas. Posiblemente el aplicativo permite al estudiante darle sentido y significado, es decir, apropiarse del concepto de regresión lineal y no lineal, por medio de la teoría, la historia y algunos ejemplos y aplicaciones de dicho concepto.

- A partir de nuestra experiencia docente y de la indagación realizada en la presente propuesta, podemos afirmar que se evidencia ausencia de la estadística en los planes de estudio de matemáticas, y en consecuencia la construcción y comprensión de los modelos de regresión en la educación media; por tal motivo los aplicativos y simulaciones están diseñados de tal manera que permitan desarrollar y abordar las competencias y estándares planteados por el MEN (2006) en el Pensamiento Aleatorio y Sistemas de datos. Es así como podemos afirmar que tiene una estructura curricular acorde con las exigencias nacionales, respecto a la educación matemática y estadística en Colombia.

- Para todo proceso de indagación realizado desde su planeación, pasando por su diseño y concluyendo en su reflexión y análisis, desarrollar este tipo de propuestas permite contribuir a nuestra formación como docentes de matemáticas y enfocar nuestro perfil como docentes en educación estadística, mostrando así en este trabajo nuestro compromiso con el conocimiento matemático y estadístico, no sólo desde una perspectiva epistémica sino también desde la posible transformación de las prácticas educativas y pedagógicas en estadística, en el contexto de la educación media.

Elaborado por:	Catalina Murcia y Fredy Peña
Revisado por:	Benjamín Sarmiento

Fecha de elaboración del Resumen:	03	12	2012

TABLA DE CONTENIDO

LISTA DE ANEXOS

INTRODUCCIÓN

Para desarrollar la presente propuesta se presentan cuatro capítulos que tienen los componentes de todo el proceso de indagación. En el primer capítulo se muestra el proceso de construcción y planificación realizado para el desarrollo del trabajo, que inicia dando una mirada a la reconstrucción teórica de los modelos de regresión lineal y no lineal, realizada a partir del estudio de investigaciones, tesis de diferentes niveles de educación profesional, artículos de revistas, entre otros; los cuales son presentados como los antecedentes de la propuesta.

Posteriormente, se encuentra la justificación del trabajo y se delimita y describe el tema de interés, del cual surgen las preguntas específicas de interés para el estudio, las cuales constituyen una pregunta global que guía el desarrollo de la investigación, y por medio de estas se plantean los objetivos específicos y general del trabajo de grado.

Seguidamente, se da a conocer un plan de acción que busca abordar el tema de interés planteado, donde se presenta la población a la que está dirigido el trabajo y especificando las fases de la indagación durante el tiempo de su realización.

En el segundo capítulo se dan a conocer los aspectos que intervienen en la construcción y caracterización del marco teórico y didáctico, el cual compone los desarrollos a través del tiempo de investigaciones a partir de tres dimensiones relacionadas con lo epistémico, didáctico y curricular, referente a los modelos de regresión lineal y no lineal.

En cuanto a la caracterización de la metodología, presentada en el capítulo tres, se realiza un recuento y explicación de los elementos que se tienen en cuenta durante el desarrollo y diseño de la propuesta. Es en este mismo capítulo, donde se realiza una descripción de los simuladores y aplicativos construidos, a través de la identificación de las partes y componentes del software, para después relacionar dichos aplicativos con los estándares en matemáticas de grado 11º referidos al objeto de estudio. Por último, en este capítulo, se dan a conocer algunos ejemplo de las actividades que se pueden abordar haciendo uso de los aplicativos y simuladores.

Finalmente se presenta el capítulo cuatro, que muestra las reflexiones y análisis en torno a les tres dimensiones trabajadas en los antecedentes y marco teórico. Además se realiza una reflexión respecto a los aportes que el trabajo brinda en general a los profesores de matemáticas y en particular a los autores de la presente propuesta.

CAPÍTULO 1. PROCESO DE INDAGACIÓN

1.1. INTRODUCCIÓN

En este capítulo se da a conocer el proceso de construcción y planificación realizado para el desarrollo de la presente propuesta. Esta descripción, inicia dando una mirada a la reconstrucción teórica hecha del tema por abordar "*la regresión lineal y no lineal"*, realizada a partir del estudio de investigaciones, tesis de diferentes niveles de educación profesional, memorias de eventos matemáticos, artículos de revistas, entre otros; y presentados como los antecedentes de este trabajo.

Partiendo de estos antecedentes, se justifica la propuesta y se delimita el tema de interés, conformando así el problema y punto de partida del estudio, que en este caso se centra en la necesidad de aportar al uso de la regresión, no sólo lineal, sino de la exploración de los diferentes tipos de regresión, en problemas. De esta problemática, surgen preguntas específicas de interés para el estudio, las cuales constituyen unas preguntas (orientadora y auxiliares) que guían el desarrollo del trabajo, y por medio de éstas se plantean los objetivos específicos y el general de la propuesta.

A partir de lo anterior, se busca un plan de acción para abordar el interés, donde se clarifica la población a la que está dirigido el estudio y la forma como se diseñará y analizará el trabajo realizado. Por último, se especifica cada momento o fases del trabajo durante el tiempo de su realización.

1.2. ANTECEDENTES

1.2.1. Dimensión Enseñanza y aprendizaje

(Godino, 1995) en su artículo *"¿qué aportan los ordenadores a la enseñanza y aprendizaje de la estadística?"* muestra un análisis de la influencia de los ordenadores en el desarrollo y expansión de la estadística en los últimos años. Por tal motivo, desarrollan un estudio centrado en la necesidad de incentivar la renovación de los contenidos y las metodologías en la enseñanza de la estadística. Es así, como hace una propuesta donde se presentan ayudas para el aprendizaje de conceptos y métodos estadísticos, haciendo uso de los ordenadores.

(Sanchez, 1999) En su tesis titulada "*Significado de la correlación y regresión para los estudiantes universitarios*", hace un estudio descriptivo del razonamiento correlacional; analizando, la presentación de la correlación y regresión en los libros de texto de bachillerato desde una mirada teórica y didáctica, así como los apuntes de clase de profesores y de dos estudiantes. A partir de esto, el autor genera un instrumento de evaluación que permite analizar conocimientos conceptuales y procedimentales, y algunas traducciones de las representaciones de correlación.

(Batanero, C. & Díaz, C., 2000) en su artículo *"El Papel de los Proyectos en la Enseñanza y Aprendizaje de la Estadística"* presenta una propuesta del trabajo en educación estadística, haciendo uso de los proyectos de aula, algunos de los cuales son planteados por el profesor y otros escogidos libremente por los alumnos. Esto lo hacen a partir de un ejemplo de proyecto, que puede ser reproducido por los profesores de secundaria o bachillerato en sus clases de estadística.

(Lopez y otros, s.f.) en su artículo "*Excel como una herramienta asequible en la enseñanza de la Estadística"*, analizan algunos de los objetivos educativos que la disponibilidad de las computadoras y paquetes informáticos, plantea sobre la enseñanza de la estadística, dentro de los cuales reconoce la importancia del manejo de las TIC en la enseñanza de la estadística.

1.2.2. Dimensión epistemológica

(Rivas, 2004) en su artículo "*Regresión no lineal*" muestra un panorama general del concepto de regresión no lineal, discutiendo la forma de estimar los parámetros en este tipo de modelos, analizando sus propiedades asintóticas y la construcción asintótica de intervalos de confianza y pruebas de hipótesis. Por último, da una aplicación con actividades haciendo uso de un modelo no lineal.

(Arias, 1986) en su Tesis de Maestría *"Regresión no lineal un enfoque práctico"*, expone tres métodos usualmente utilizados en la regresión no lineal, con los cuales presenta algoritmos computacionales referidos a estos métodos de regresión, ya que el indica que la solución de un problema práctico requiere resolverlo computacionalmente. Finalmente reflexiona acerca de la importancia de la solución de problemas cotidianos y del uso de los computadores en el manejo de la regresión lineal.

(Rojas, 1990) en su artículo *"Técnicas de diagnóstico en Regresión Lineal*", presenta diferentes técnicas que permiten detectar observaciones relacionadas con la regresión lineal; es así como se hace un estudio comparativo y se genera un método adecuado y pertinente para disminuir el efecto de los datos atípicos.

1.2.3. Dimensión Curricular

El Ministerio Educación Nacional MEN (1998) en los *"Lineamientos Curriculares para el Área de Matemáticas"*, presenta un conglomerado de orientaciones para el área de matemáticas de la educación básica y media, las cuales son clasificados en cinco pensamientos (numérico y sistemas numéricos, espacial y sistemas geométricos, métrico y sistemas de medidas, aleatorio y los sistemas de datos, variacional y sistemas algebraicos y analíticos).

El Ministerio Educación Nacional MEN (2006) en los *"Estándares Básicos del Calidad en Matemáticas"*, propone unos criterios claros, que son el punto de referencia de lo que el estudiante puede estar en capacidad de *hacer* y *saber hacer*, a partir de cinco campos de pensamientos (el numérico, el espacial, el métrico o de medida, el variacional y el aleatorio o probabilístico) , donde se pretende abordar todas la competencias, capacidades y aptitudes en el área de matemáticas. Cada uno de los cinco pensamientos son profundizados mediante niveles de curso que están determinados en grupos (1° a 3°, 4° a 5°, 6° a 7°, 8° a 9° y 10° a 11°).

Los anteriores lineamientos y estándares para el área de matemáticas y en particular para la enseñanza de la estadística son abordados y explicados en el presente trabajo, en el capìtulo 2 "*Dimensión Curricular*".

1.3. JUSTIFICACIÓN

En busca de diseñar estrategias y modelos que incentiven a la construcción y mejoramiento de los planes de estudio de la Educación Matemática y Estadística de grado 11º, se plantean propuestas metodológicas de innovación, que involucren las TIC, para la enseñanza de estas ciencias.

Dichas estrategias, según Batanero (2001), han aumentado los contenidos estadísticos a enseñar, donde se le da mayor importancia a los aspectos interpretativos y conceptuales y menor a los procedimentales y algoritmos de cálculo, "que permite a los estudiantes experimentar y explorar todos los aspectos de los procesos estadísticos, desde la planificación de la muestra o del diseño experimental hasta la recolección y el manejo de datos, la simulación y el análisis, para interpretar y comunicar los resultados, (…) es así como se enriquece el significado de los conceptos mostrados a los estudiantes." (Batanero, 2001, p. 36)

De la misma manera se dirige la mirada hacia la cultura estadística (Gal, 2000, citado por Batanero, 2001, p. 36) e informatizada donde se debe promover "una comprensión de las técnicas básicas de análisis de datos y de su interpretación, que son cada día más importantes. Esto nos lleva a tener que enseñar estadística a alumnos con capacidades y actitudes variables (...) donde los estudiantes lleguen a comprender y a apreciar el papel de la estadística en la sociedad, conociendo sus diferentes campos de aplicación y el modo en que la estadística ha contribuido a su desarrollo". (Hoyos y otros, 2006, citado por Batanero, 2001, p. 38)

Por esta razón, vemos la necesidad de aportar al planteamiento de nuevas iniciativas que generen el fortalecimiento de conceptos estadísticos relacionados con la correlación y regresión lineal y no lineal, correspondientes al objeto de estudio, propiciando alternativas a nivel pedagógico y didáctico, y apoyando los procesos de enseñanza y aprendizaje de la educación matemática y estadística.

En consecuencia, se pensó en el desarrollo de un material educativo como agente mediador en el proceso de enseñanza y aprendizaje del objeto de estudio. Será entonces, un mecanismo a través del cual el estudiante tendrá las herramientas necesarias para identificar, caracterizar, analizar y utilizar la correlación y los modelos de regresión lineal y no lineal en contextos reales, donde, de acuerdo con Fernandez, Soler y Sarmiento (2008), es necesario que los estudiantes aprendan los contextos en los cuales es aplicable o no un procedimiento o concepto enseñado. Es decir, que identifiquen el modelo de regresión que deben utilizar de acuerdo con el tipo de situación presentada.

Este material permitirá al estudiante determinar algunos aspectos relacionados con el objeto de estudio, donde no sólo hará uso de los procedimientos y algoritmos matemáticos y estadísticos, sino que centrará su mirada en el análisis y aplicación de los conceptos relacionados con la regresión lineal y no lineal. Es así como, según Fernandez, Soler, y Sarmiento (2008), no tiene sentido que los alumnos procedan a la repetición de cálculos tediosos para intentar aumentar su destreza de cálculo, sino que se debe hacer énfasis en la importancia que tiene la interpretación, el análisis y la resolución de problemas cotidianos estadísticos.

De la misma manera, el profesor tendrá la oportunidad de contar con un recurso electrónico que facilite la enseñanza de los conceptos relacionados con el objeto de estudio, que contiene situaciones que se ubican en algún contexto de aplicación. Es así, como los aportes que se logren con el diseño de este material educativo podrá ser una experiencia que enriquezca y fortalezca la enseñanza matemática y estadística.

1.4. DESCRIPCIÓN DEL INTERÉS

En los últimos años la estadística ha ido adhiriéndose a los planes de estudio de matemáticas en Colombia, dado que existen, según Holmes (1980, citado por, Batanero C. , s.f.), algunos intereses hacia la enseñanza de la misma, dentro de los cuales se encuentra: la perspectiva de la estadística como una parte de la educación importante, pertinente e indispensable para la vida; la utilidad que tiene esta rama de las matemáticas dentro de las distintas profesiones; y la ayuda que le brinda a la comprensión de otros temas y competencias del currículo de matemáticas.

A partir de esto, se genera la necesidad de buscar estrategias que permitan abordar el proceso de enseñanza – aprendizaje de la estadística, para lo cual varios autores han propuesto el uso de ordenadores como una herramienta para motivar y ayudar a comprender y construir los conceptos estadísticos con facilidad. Un ejemplo de esto, es el trabajo realizado por Godino (1995) quien indica que “la capacitación estadística incluye hoy día el conocimiento del modo de procesar datos mediante un programa estadístico, por lo que deberíamos, en la medida de lo posible, ofrecer a nuestros alumnos un primer contacto con este tipo de programas. Por otro lado, el ordenador no es sólo un recurso de cálculo, sino un potente útil didáctico, que nos permite conseguir una aproximación más exploratoria y significativa en la enseñanza de la estadística”.

En consecuencia se piensa en plantear una propuesta a un tema estadístico que, desde lo evidenciado en nuestras experiencias, no haya sido trabajado en grado 11º, ni haya sido abordado en los planes de estudio de la educación media. Dadas estas características, se escoge los *modelos de regresión lineal y no lineal* como objeto de estudio para el desarrollo de la presente propuesta

Por esta razón y en busca de incluir nuevas herramientas y técnicas que apoyen los procesos de enseñanza y aprendizaje, y que permitan la comprensión y el fortalecimiento de conceptos relacionados con la regresión lineal y no lineal, en grado 11, y que además genere en estos estudiantes la necesidad de explorar los distintos tipos de regresión para una situación cotidiana, se propone el diseño de un aplicativo que contenga los elementos suficientes para el trabajo de estas temáticas en el aula de clases de matemáticas.

1.5. PREGUNTA DE INVESTIGACIÓN

¿Cómo ilustrar y comparar los diferentes modelos de regresión lineal y no lineal para que sean explorados por estudiantes de grado 11°?

1.5.1. Preguntas auxiliares

¿Qué tipo de situaciones, problemas o ejercicios proponer para generar la exploración de la regresión lineal y no lineal en grado 11°?

¿A través de qué programa y de qué manera se pueden diseñar aplicativos y simuladores de los modelos de regresión lineal y no lineal?

¿Qué se puede concluir frente al uso de los simuladores y aplicativos en el manejo de los modelos de regresión lineal y no lineal?

1.6. OBJETIVOS

1.6.1. Objetivo General

Diseñar simulaciones y aplicativos para ilustrar y comparar, a partir de problemas, los diferentes modelos de regresión lineal y no lineal abordados en grado 11.

1.6.2. Objetivos específicos

- Proponer actividades que permitan explorar los diferentes modelos de regresión lineal y no lineal para realizar pronósticos.

- Diseñar aplicativos y simulaciones de los modelos de regresión lineal y no lineal, haciendo uso del programa Descartes.

- Reflexionar acerca de la importancia y pertinencia de los simuladores y aplicativos en el proceso de enseñanza – aprendizaje de los modelos de regresión lineal y no lineal.

1.7. PLAN DE ACCIÓN

La propuesta se lleva a cabo en cuatro fases, dando inicio con la identificación del interés y concluyendo con el análisis e informe final, se espera que durante el desarrollo de esta se abarque el problema, contando con un periodo de un año, aproximadamente. A continuación se presenta una descripción que contiene las fases y lo que se pretende hacer en cada una de ellas.

1.7.1. Fases del método de indagación

1.7.1.1. Primera Fase. Identificación del interés y proceso de indagación.

Se logra a partir de la evidencia de factores y situaciones que influyen en la necesidad de aportar al uso de la regresión, no sólo lineal, sino de la exploración de los diferentes tipos de regresión, en la solución de problemas. Además, se realiza con el fin de abordar diferentes ambientes y contextos relacionados con la educación estadística de grado 11°, que apuntan a intereses propios de los autores de la presenta propuesta.

Interés base: A partir de las experiencias propias en el aula y de los puntos de vista de algunos autores que han trabajado en la enseñanza de la estadística, surge un interés base que apunta a la necesidad de incluir herramientas y técnicas que permitan la comprensión y fortalecimiento de la regresión lineal y no lineal, manejada en grado 11.

Búsqueda de antecedentes: Luego de la identificación del problema, de la población señalada, se realizó una búsqueda y recopilación de trabajos, propuestas de investigación, tesis y demás documentos que dieran un panorama más objetivo, con el fin de sustentar y dar argumentos a la presente propuesta. A partir de esta búsqueda de antecedentes, se determinó que una posible solución a la problemática planteada es el diseño de un aplicativo que permita la exploración de los diferentes tipos e regresión Dichos antecedentes son clasificados en tres grandes dimensiones:

- **Dimensión Enseñanza y aprendizaje:** Comprende documentos referidos a la didáctica empleada para la construcción de conceptos relacionados con los modelos de regresión lineal y no lineal, así como la dificultades del aprendizaje de dicho objeto estadístico. También hace referencia al uso de las TIC en la enseñanza del objeto de estudio.

  Búsqueda acerca de herramientas de enseñanza de los modelos de regresión: En esta parte se hizo énfasis en la búsqueda de trabajos relacionados con el diseño de aplicativos para los distintos modelos de regresión, los cuales ampliaron el horizonte para afrontar e iniciar el proceso de solución de la problemática planteada.

- **Dimensión Epistémica:** Se refiere al saber, al objeto de estudio correspondiente a la regresión lineal y no lineal. Se centra en los conceptos y realiza una pequeña reseña histórica de los objetos estadísticos que allí intervienen.

- **Dimensión Curricular:** Comprende aspectos relacionados con las políticas educativas de nuestro país, en materia de educación estadística.

1.7.1.2. Segunda Fase. Observación del entorno educativo.

En esta fase se pretende obtener información pertinente y concreta para el desarrollo de la propuesta, con el fin de realizar la caracterización de los estándares básicos bajo los cuales se rigen los currículos de grado 11° en nuestro país. A partir de las observaciones realizadas, se plantea una propuesta contextual que responda a las necesidades y exigencias de los estamentos nacionales a nivel de educación matemática y estadística.

1.7.1.3. Tercera Fase. Diseño de simuladores y aplicativos

En esta fase se inicia el trabajo de exploración del programa Descartes, para realizar el diseño y montaje de los aplicativos relacionados con la regresión lineal y no lineal. El diseño tanto de los aplicativos como de las situaciones allí propuestas son elaborados a partir de los estándares y las políticas nacionales planteadas para el grado 11.

1.7.1.4. Cuarta Fase. Reflexiones

En esta fase se realiza una reflexión acerca de la importancia y pertinencia del uso de los aplicativos para la enseñanza y aprendizaje de la regresión lineal y no lineal. Además, se hace una reflexión frente a los aprendizajes propios adquiridos, durante el desarrollo y construcción de la presente propuesta.

1.7.2. Cronograma

A continuación se presenta el cronograma de las fases anteriormente mencionadas.

Fase	Feb	Mar	Abr	May	Jun	Jul	Ago	Sept	Oct
Primera Fase. Identificación del interés y proceso de indagación									
Segunda Fase. Observación del entorno educativo									
Tercera Fase. Diseño de simuladores y aplicativos									
Cuarta Fase. Análisis y Reflexiones									

CAPÍTULO 2. REFERENTES TEÓRICOS Y DIDÁCTICOS

2.1. I NTRODUCCIÓN

Dado que el proceso de indagación debe estar sustentado a partir de un marco de referentes teóricos y didácticos para su validación, en este capítulo se dan a conocer todos los aspectos tenidos en cuenta en su construcción y caracterización. Éste compone los desarrollos a través del tiempo de investigaciones a nivel teórico y didáctico referente a los conceptos relacionados con la regresión lineal y no lineal.

Las perspectivas teórica y didáctica, se constituyen dentro de este trabajo como referentes del objeto matemático, y de las metodologías y estrategias empleadas para abordarlo. En el primero, se definen los desarrollos teóricos de la regresión lineal y no lineal, estudiando sus definiciones, características y desarrollos históricos. Y, el segundo, da a conocer las distintas formas de abordar el conocimiento estadístico, en particular el concepto de la regresión, visto desde dos aspectos: desde el proceso de enseñanza y aprendizaje y desde las políticas educativas de matemáticas que actualmente se rigen en nuestro país.

2.2. DIMENSIONES TEÓRICAS Y CONCEPTUALES

2.2.1. Dimensión epistémica

En esta sección exponemos los referentes matemáticos que fundamentan todo el estudio; en este caso dichos referentes giran en torno a temáticas tales como correlación, regresión, método de los mínimos cuadrados y regresión de tipo lineal y no lineal. Primeramente presentamos un breve recuento histórico de los hechos que dieron origen a tales temas, seguido por una descripción de los mismos.

2.2.1.1. Reseña histórica

A continuación se presentan los referentes históricos relacionados con regresión, método de los mínimos cuadrados y coeficiente de correlación.

Hacia el año 1801, específicamente el 1 de enero, el sacerdote, educador y astrónomo Giuseppe Piazzi descubrió desde un observatorio en Palermo (Italia) el asteroide Ceres, luego denotado como planeta enano por la Unión Astronómica Internacional, Piazzi logró seguir la trayectoria de Ceres a lo largo de 40 días, luego de este tiempo no pudo volver a encontrarlo, por lo que se convirtió en un reto para los sabios de la época el predecir en qué lugar debía reaparecer el cuerpo celeste, pero el único método conocido en ese entonces para el problema se remitía a dar solución a las ecuaciones no lineales de Kepler, lo cual era muy difícil. Fue entonces cuando Carl Friedrich Gauss encontró la solución tan buscada por medio de un método bastante sencillo denominado el **método de los mínimos cuadrados,** publicado solo hasta el año 1809 en el segundo volumen del libro *Theoria*

Motus Corporum Coelestium in sectionibus conicis solem ambientium, por este descubrimiento se le otorgó a Gauss el nombramiento como director del Observatorio Astronómico de Göttingen desde 1807 hasta el día de su muerte.

El método se constituye en el primer estudio y además el más común para ajustar una función lineal a un conjunto de parejas ordenadas, de las que se presume existe una correlación o dependencia lineal de una variable (independiente) sobre la otra (dependiente), este proceso se conoce con el nombre de regresión lineal.

Por otro lado, el término correlación, referido al grado de relación o dependencia entre dos variables surge a mediados del siglo XIX, dada la necesidad de estudiar variables posiblemente relacionadas entre si, particularmente en el campo de la medicina. El coeficiente de correlación lineal, fue propuesto por primera vez el médico inglés Sir Francis Galton, nacido en el año 1822, quien, debido a sus estudios sobre genética propuso hacia el año 1869 en el libro “Hereditary Genius” un número que permitía establecer el grado de dependencia lineal que tienen un par de variables cuantitativas.

Gracias a los estudios de Galton, el matemático Karl Pearson, en octubre de 1894 comenzó a promover en sus estudiantes del University College de Londres, el uso de la estadística, que hasta el momento era únicamente descriptiva, hacia un enfoque investigativo, usando la descripción de los datos y la dependencia de una variable sobre la otra, para predecir posibles comportamientos de una variable, al modificar la otra, este trabajo, le fue suficiente a Pearson, para aseverar que, dado que todas las inferencias se basan entre la relación entre unos antecedentes y unos consecuentes, entonces el razonamiento científico posee una base netamente estadística.

2.2.1.2. Reconstrucción matemática

Correlación:

La correlación en estadística es un mecanismo que permite determinar si un par de variables cuantitativas aleatorias guardan o no algún tipo de relación o dependencia, por ejemplo, en una situación de caída libre, las variables tiempo y velocidad están directamente relacionadas, de hecho responden a un modelo matemático estandarizado que ajusta de la mejor manera los resultados que puedan obtenerse en experimentos reales.

Para saber que tan relacionados linealmente se encuentran un par de variables X e Y, debemos calcular las desviaciones típicas de las variables de manera independiente, así:

$$\delta_X = \sqrt{\sum_{i=1}^{n} \frac{x_i^2}{n} - \overline{X}^2}$$

$$\delta_Y = \sqrt{\sum_{i=1}^{n} \frac{y_i^2}{n} - \overline{Y}^2}$$

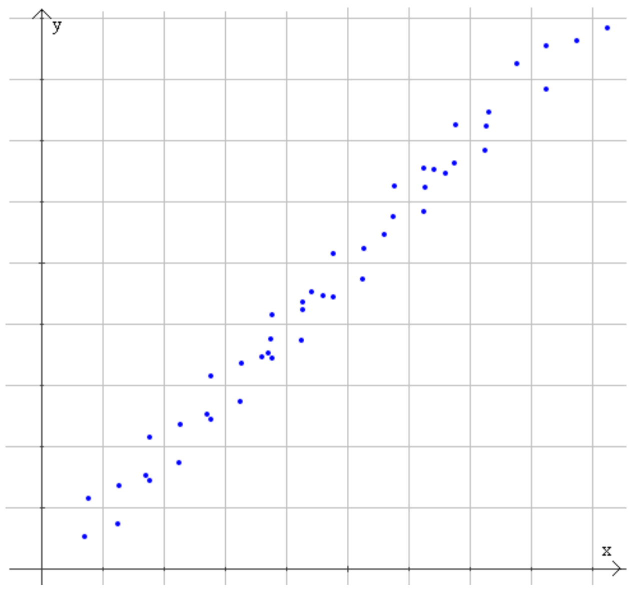
Calculamos además la covarianza definida de la siguiente manera.

$$\delta_{XY} = \sum_{i=1}^{n} \frac{x_i y_i}{n} - \overline{X}\,\overline{Y}$$

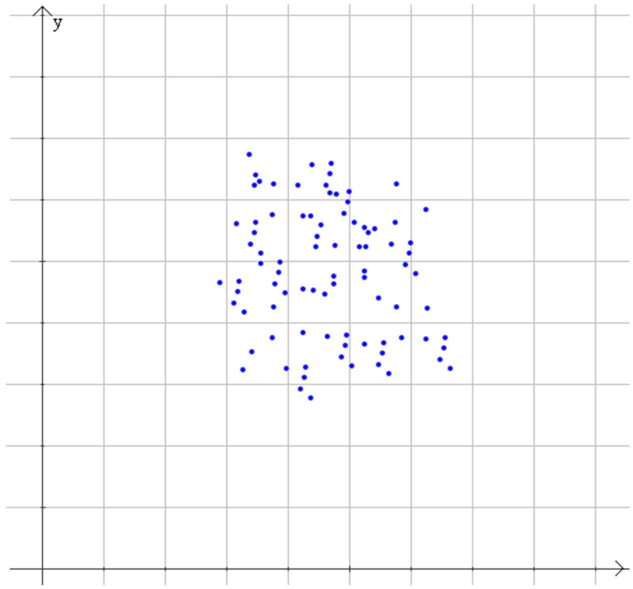
Con la covarianza, determinamos un coeficiente de correlación con la siguiente ecuación:

$$r = \frac{\delta_{XY}}{\delta_X \delta_Y}$$

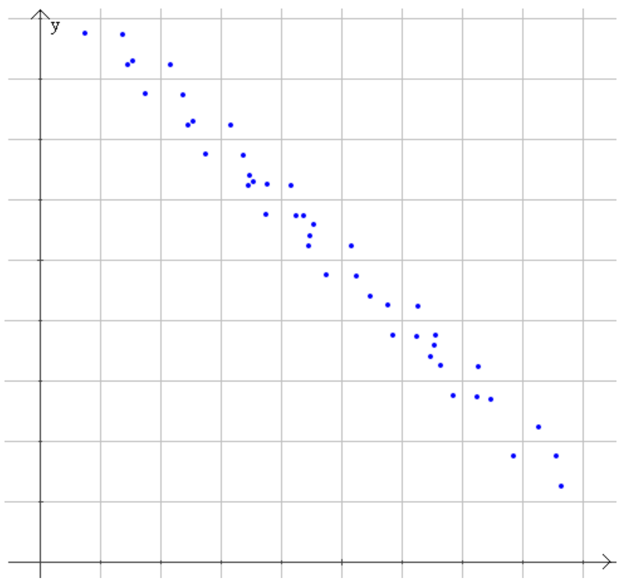
Los resultados de r se encuentran en el intervalo $[-1,1]$, al obtener $r = 1$ significa que existe una correlación perfecta, es decir que una de las variables depende de la otra de manera directa y lineal, si $0 < r < 1$ significa que existe correlación positiva es decir que se ajusta a un modelo de la forma $Y = mX + \beta$, con $m > 0$ como se muestra en la siguiente figura.



Si $r = 0$ no existe correlación lineal, aunque las variables pueden estar relacionadas de manera no lineal, un ejemplo puede ser el de la siguiente figura para el cual no hay correlación de ningún tipo, ya que no existe una relación entre una variable y la otra y por tanto no se evidencia un comportamiento determinado de la nube de puntos:



Si $r = -1$ entonces existe una correlación negativa perfecta, es decir que los datos reales corresponden a los datos que se obtienen del modelo; si $-1 < r < 0$ existe una correlación negativa es decir que se ajusta a un modelo de la forma $Y = mX + \beta$, con $m < 0$, como se muestra en la siguiente figura:



Coeficiente de correlación generalizado

Dado que el coeficiente de correlación lineal se aplica únicamente cuando se presume que los datos se encuentran relacionados de manera lineal, un coeficiente de correlación general para cualquier tipo de regresión puede definirse como el cociente entre la variación total de y con la variación esperada para esa variable, de la siguiente manera:

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\sum_{i=1}^{n}(y_{est} - \bar{y})^2}$$

Donde y_{est} corresponde al valor estimado de y para un valor dado de x, obtenido de la curva de regresión de y sobre x.

Esta última definición refleja la forma de la curva de la regresión y de este modo es apropiada como la definición de un coeficiente de correlación generalizado r, el cual se ajusta a coeficientes de correlación no lineales, donde se mide qué tan bien se ajusta una curva de regresión no lineal a los datos.

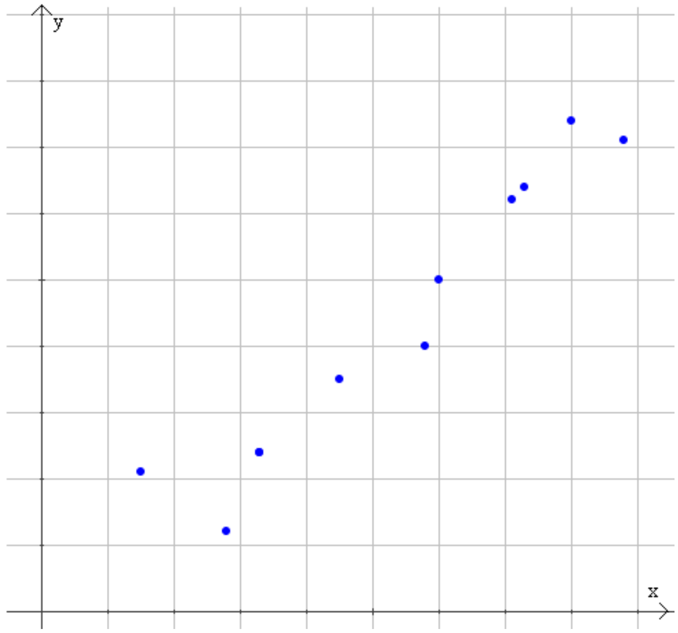
Regresión

De acuerdo con la disposición que tengan los datos una vez representados en el plano cartesiano se espera establecer un modelo adecuado, por tanto existen una gran variedad de modelos de regresión, entre ellos se encuentran:

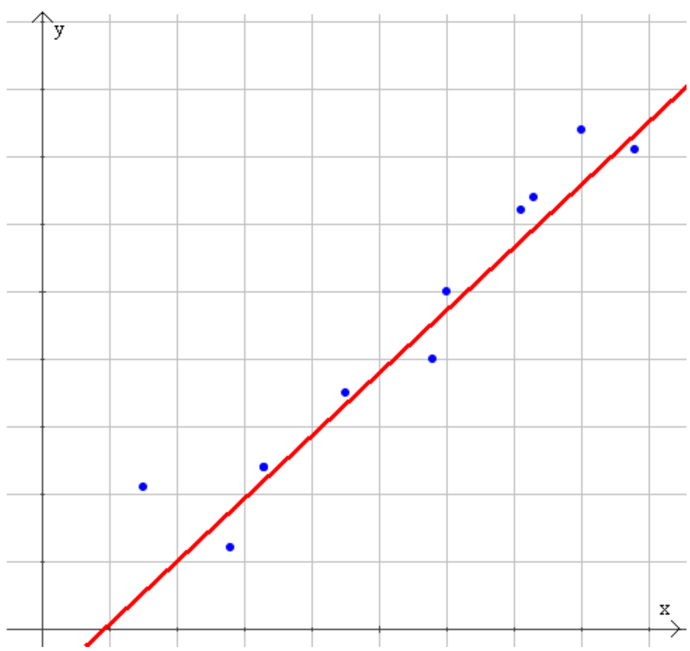
- Modelo de regresión lineal
- Modelo de regresión polinómica
- Modelo de regresión cuadrática
- Modelo de regresión logística
- Modelo de regresión exponencial

Modelo de regresión lineal

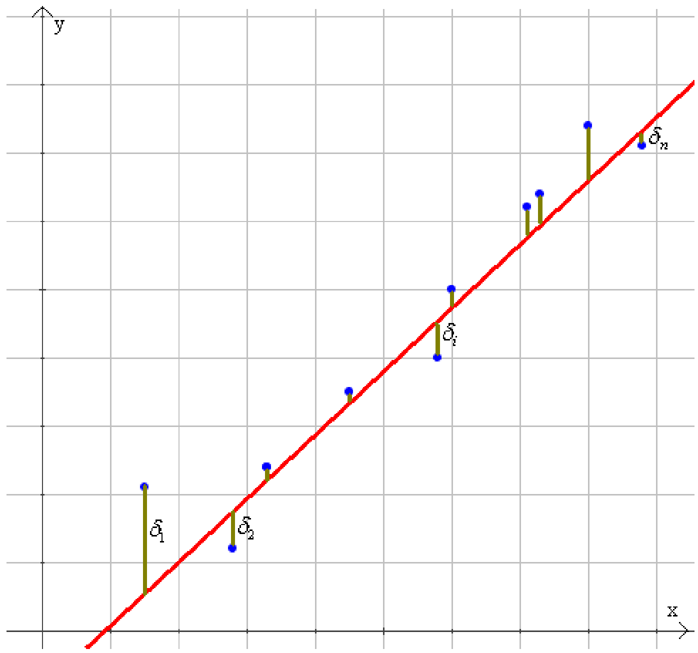
En primera medida el modelo de regresión lineal otorga la ecuación de una recta en el plano, explicaremos este modelo por medio del **método de los mínimos cuadrados.**
Supongamos un conjunto de datos (x_i, y_i), con una cantidad de elementos n, dado que son parejas ordenadas pueden representarse en un plano cartesiano de esta manera:



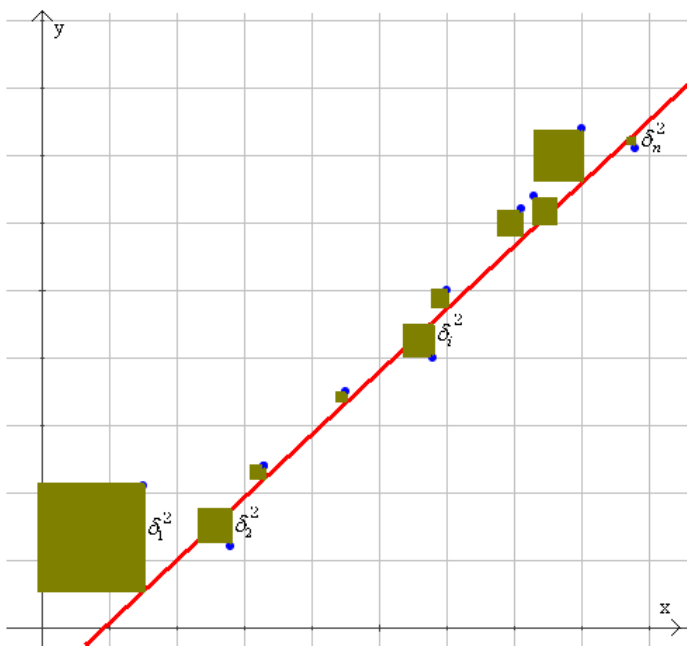
Se pretende obtener una recta que se ajuste de la mejor manera a la dispersión de puntos que se encuentran**.** De manera que se obtendría un gráfico como el siguiente.



Teniendo en cuenta que queremos obtener el modelo lineal que mejor se ajuste, buscamos las diferencias entre las ordenadas de los puntos y la función de regresión calculada en la abscisa correspondiente, ésta también se conoce como la distancia vertical, de la siguiente manera:



Elevamos al cuadrado cada una de estas diferencias.



Definimos el error total como la suma de los cuadrados de tales distancias, es decir:

$$E=\sum_{i=1}^{n}\left(\delta_i\right)^2$$

Dado que la función lineal debe tener la forma $y=mx+\beta$ el error sería:

$$E=\sum_{i=1}^{n}\left(y_i-mx_i-\beta\right)^2$$

Necesitamos encontrar los valores de m y β que minimicen tal error, para ello utilizamos herramientas del cálculo diferencial, es decir vamos a calcular las derivadas parciales de la función de error respecto a m y a β

$$\frac{\partial E}{\partial m}=-2\sum_{i=1}^{n}x_i\left(y_i-mx_i-\beta\right)$$

$$\frac{\partial E}{\partial \beta}=-2\sum_{i=1}^{n}\left(y_i-mx_i-\beta\right)$$

Teniendo presente que debemos optimizar el error, sus derivadas parciales deben igualarse a cero de la siguiente manera

$$0=-2\sum_{i=1}^{n}x_i\left(y_i-mx_i-\beta\right)$$

$$0=-2\sum_{i=1}^{n}\left(y_i-mx_i-\beta\right)$$

O escrito de otra forma:

$$0=\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i^{\,2}m-\sum_{i=1}^{n}x_i\beta$$

$$0=\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}mx_i-n\beta$$

Entonces obtenemos un sistema de ecuaciones de 2×2 de esta manera denominadas ecuaciones normales de regresión:

$$1)\quad m\sum_{i=1}^{n}x_i^{\,2}+\beta\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_iy_i$$

$$2)\quad m\sum_{i=1}^{n}x_i+n\beta=\sum_{i=1}^{n}y_i$$

Si despejamos de 2) la incógnita β obtenemos:

$$\beta = \frac{\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i}{n}$$

Sustituyendo esto en la primera ecuación se obtiene:

$$m\sum_{i=1}^{n} x_i^{\,2} + \frac{\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i}{n}\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

O escrito de otra forma:

$$\frac{m\left(n\sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2\right) + \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i}{n} = \sum_{i=1}^{n} x_i y_i$$

De manera que despejando *m*obtenemos:

$$m = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2}$$

De manera pues que obtenemos las formulas correspondientes para determinar m y β .

$$\boxed{m = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2}}$$

$$\boxed{\beta = \frac{\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i}{n}}$$

Con estas ecuaciones se puede demostrar que la función de la recta $y = mx + \beta$ se cumple para la pareja $\left(\overline{x}, \overline{y}\right)$ es decir para los valores medios de las variables. Dicha demostración se hará a continuación:

Demostración: Debemos probar que $\bar{y} = m\bar{x} + \beta$, demostraremos esto, partiendo de la segunda de las ecuaciones normales de regresión:

$$m\sum_{i=1}^{n} x_i + n\beta = \sum_{i=1}^{n} y_i$$

Al dividir por *n* cada lado de la igualdad se obtiene:

$$m\frac{\sum_{i=1}^{n} x_i}{n} + \beta = \frac{\sum_{i=1}^{n} y_i}{n}$$

Es decir: $m\bar{x} + \beta = \bar{y}$

Queda entonces demostrado que la los valores medio de las variables pueden ser expresado como una ecuación lineal, en otras palabras, se puede decir que lafunción de regresión lineal se cumple para los valores medio de las variables.

Ejemplo regresión lineal

El Instituto Nacional de Medicina Legal y Ciencias Forenses, recoge y analiza la información sobre lesiones fatales y no fatales de causa externa para darlas a conocer y colaborar con la formulación y seguimiento a las políticas públicas diseñadas con el fin de reducir el fenómeno de la violencia en Colombia.

Es así como a través de los puntos de atención en donde el INMLCF tiene presencia (información directa) y por medio de los médicos rurales en los lugares del país en donde no se cuenta con unidad básica (información indirecta), se ingresa la información que permite caracterizar las lesiones de causa externa mediante variables sociodemográficas, de caracterización del hecho y espacio-temporales.

A continuación se presenta, de manera detallada, la información estadística de los homicidios durante los años 2004 a 2009 de la ciudad de Bogotá:

Departamento del hecho	2004	2005	2006	2007	2008	2009	Total
Bogotá, D.C.	1600	1689	1336	1401	1465	1649	9140

Teniendo en cuenta el número de homicidios en Bogotá entre los años 2004 y 2009, determine:
El coeficiente de correlación

En el año 2012, ¿cuántos homicidios podemos suponer que se presentarán en la ciudad de Bogotá?

Solución: Antes de calcular el coeficiente de correlación, Determinemos los estadísticos básicos para $\boldsymbol{x}$ e $\boldsymbol{y}$, construyendo una tabla donde determinaremos los valores cuadrados de las dos variables y el producto de las mismas, así:

$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum y_i^2$	$\sum x_i y_i$
2004	1600	4016016	2560000	3206400
2005	1689	4020025	2852721	3386445
2006	1336	4024036	1784896	2680016
2007	1401	4028049	1962801	2811807
2008	1465	4032064	2146225	2941720
2009	1649	4036081	2719201	3312841
12039	**9140**	**24156271**	**14025844**	**18339229**

Determinemos la media para las variables ***x*** e ***y***:

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{12039}{6} = 2006.5$$

$$\bar{y} = \sum_{i=1}^{n} \frac{y_i}{n} = \frac{9140}{6} = 1523.33$$

Ahora, a partir de los valores hallados determinemos las desviaciones típicas de ***x*** y***y:***

$$\delta_x = \sqrt{\sum_{i=1}^{n} \frac{x_i^2}{n} - \bar{x}^2} = \sqrt{\frac{24156271}{6} - (2006.5)^2} = 2.91$$

$$\delta_y = \sqrt{\sum_{i=1}^{n} \frac{y_i^2}{n} - \bar{y}^2} = \sqrt{\frac{14025844}{6} - (1523.33)^2} = 17096.22$$

Teniendo las desviaciones típicas de las dos variables, calculemos entonces las covarianza.

$$\delta_{xy} = \sum_{i=1}^{n} \frac{x_i y_i}{n} - \bar{x}\bar{y} = \frac{18339229}{6} - (2006.5)(1523.33) = -30.16$$

Entonces, podemos calcular lo que nos indican en el primer ítem, el coeficiente de correlación. Para tal fin utilizaremos la siguiente ecuación:

$$r = \frac{\delta_{xy}}{\delta_x \delta_y} = \frac{-30.16}{(2.91)(17096.22)} = \frac{-30.16}{49863.98} = 0.000604$$

Para determinar la cantidad de que se presentarán en la ciudad de Bogotá en el año 2012, debemos determinar la ecuación de la recta de regresión. Para esto, utilicemos el método de mínimos cuadrados:

Tenemos que la ecuación de la recta es $y = mx + \beta$. Con este método hallaremos los valores de *m* y *b*.

Hallemos primero la pendiente utilizando la ecuación:

$$m = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^{2} - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Luego,

$$m = \frac{6(18339229) - (12039)(9140)}{6(24156271) - (12039)^2} = \frac{110035374 - 110036460}{144937626 - 1461681} = \frac{-1086}{143475945}$$

$$m = 7.56 \times 10^{-6}$$

Ahora hallemos el punto de intersección con eje *y*:

Dado que, $\beta = \dfrac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$, tenemos

$$\beta = \frac{(9140)(24156271) - (12039)(18339229)}{6(24156271) - (12039)^2} = \frac{220788316940 - 220785977931}{144937626 - 144937521}$$

$$\beta = \frac{2339009}{105}$$

$$\beta = 2276,27$$

Teniendo el valor de la pendiente $\boldsymbol{m = 7.56\mathrm{x}10^{-6}}$y el valor del punto de corte de la recta con el eje y $\boldsymbol{\beta = 22276,27}$, reemplacemos estas dos cantidades en la ecuación de la recta:

$$y = mx + \beta$$

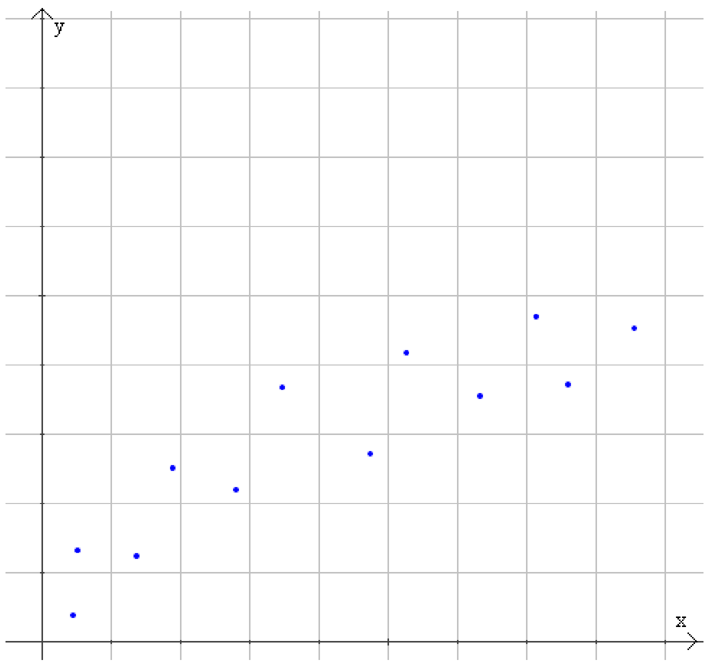
$$y = 7.56x10^{-6}(2012) + 2276,27 = 2276$$

Ahora, cantidad de que se presentarán en la ciudad de Bogotá en el año 2012, debemos reemplazar el valor 2012 en la variable x, así:

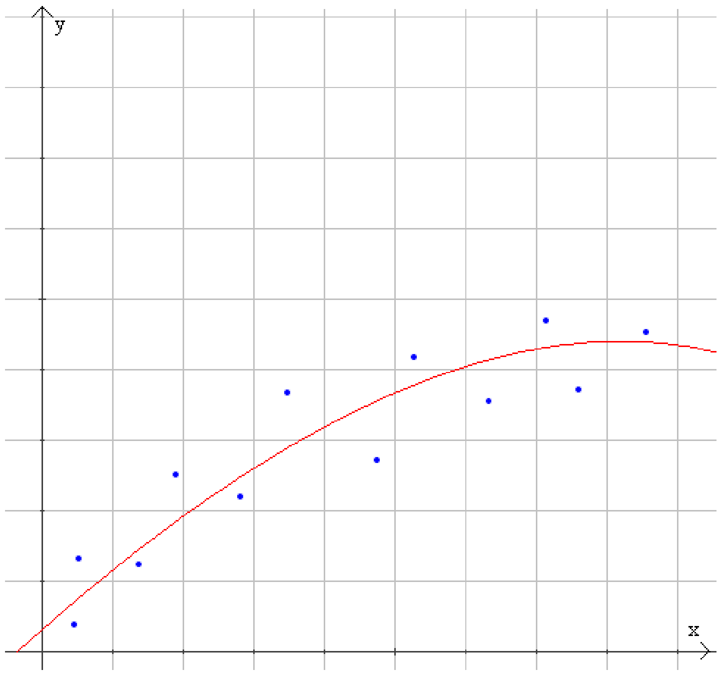
Si $x = 2012$ Entonces podemos suponer que en el año 2012 habrán 2276 homicidios en la ciudad de Bogotá.

Modelo de regresión Cuadrática

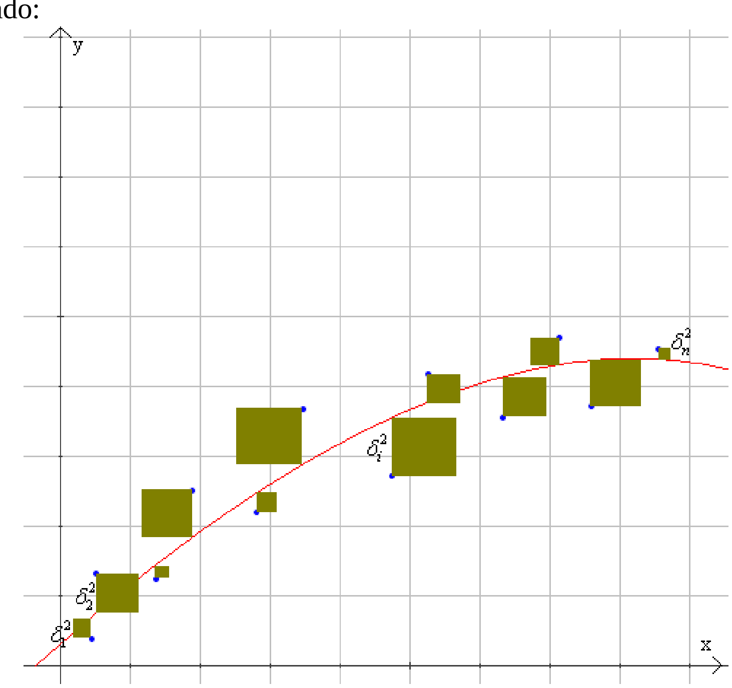
Para éste caso, de igual manera podemos hacer uso del método de los mínimos cuadrados para obtener el modelo que mejor se ajuste a la dispersión de puntos. Supongamos un conjunto de puntos (x_i, y_i) de n elementos, representados en el plano cartesiano de la siguiente manera:



Una curva cuadrática que se ajustaría mejor al conjunto de datos se representaría de la siguiente manera:



Para éste caso en particular un modelo lineal no se ajusta de una manera conveniente, dado que los puntos muestran un comportamiento curvo. Se procede a calcular las distancias verticales y elevarlas al cuadrado:



El error total de igual manera se define como la suma de los cuadrados de las distancias verticales entre el modelo y los datos, de manera que como se busca una función de la forma *f(x)*= $ax^2 + bx + c$ el error total se define como:

$$E=\sum_{i=1}^{n}\left(y_i-ax_i^2-bx_i-c\right)^2$$

Derivando respecto a los parámetros a, b y c se obtiene:

$$\frac{\partial E}{\partial a}=-2\sum_{i=1}^{n}\left(y_i-ax_i^2-bx_i-c\right)x_i^2$$
$$\frac{\partial E}{\partial b}=-2\sum_{i=1}^{n}\left(y_i-ax_i^2-bx_i-c\right)x_i$$
$$\frac{\partial E}{\partial c}=-2\sum_{i=1}^{n}\left(y_i-ax_i^2-bx_i-c\right)$$

Igualando a cero para optimizar la función de error se obtiene el siguiente sistema de ecuaciones 3×3:

$$1)\quad a\sum_{i=1}^{n}x_i^4+b\sum_{i=1}^{n}x_i^3+c\sum_{i=1}^{n}x_i^2=\sum_{i=1}^{n}y_ix_i^2$$
$$2)\quad a\sum_{i=1}^{n}x_i^3+b\sum_{i=1}^{n}x_i^2+c\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}y_ix_i$$
$$3)\quad a\sum_{i=1}^{n}x_i^2+b\sum_{i=1}^{n}x_i+cn=\sum_{i=1}^{n}y_i$$

Las soluciones del sistema de ecuaciones son:

$$a=\frac{\begin{vmatrix}\sum_{i=1}^{n}y_ix_i^2 & \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2\\ \sum_{i=1}^{n}y_ix_i & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}y_i & \sum_{i=1}^{n}x_i & n\end{vmatrix}}{\begin{vmatrix}\sum_{i=1}^{n}x_i^4 & \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2\\ \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i & n\end{vmatrix}}\quad b=\frac{\begin{vmatrix}\sum_{i=1}^{n}x_i^4 & \sum_{i=1}^{n}y_ix_i^2 & \sum_{i=1}^{n}x_i^2\\ \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}y_ix_i & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}y_i & n\end{vmatrix}}{\begin{vmatrix}\sum_{i=1}^{n}x_i^4 & \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2\\ \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i & n\end{vmatrix}}\quad c=\frac{\begin{vmatrix}\sum_{i=1}^{n}x_i^4 & \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}y_ix_i^2\\ \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}y_ix_i\\ \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i & \sum_{i=1}^{n}y_i\end{vmatrix}}{\begin{vmatrix}\sum_{i=1}^{n}x_i^4 & \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2\\ \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i & n\end{vmatrix}}$$

Unas fórmulas más simplificadas se obtienen al usar el parámetro a para determinar b y finalmente los parámetros a y b para hallar el valor de c de manera que tenemos para el modelo cuadrático las siguientes fórmulas:

$$a=\frac{\begin{vmatrix}\sum_{i=1}^{n}y_ix_i^2 & \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2\\ \sum_{i=1}^{n}y_ix_i & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}y_i & \sum_{i=1}^{n}x_i & n\end{vmatrix}}{\begin{vmatrix}\sum_{i=1}^{n}x_i^4 & \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2\\ \sum_{i=1}^{n}x_i^3 & \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i\\ \sum_{i=1}^{n}x_i^2 & \sum_{i=1}^{n}x_i & n\end{vmatrix}}$$

$$b=\frac{\sum_{i=1}^{n}x_i\left(\sum_{i=1}^{n}y_i-a\sum_{i=1}^{n}x_i^2\right)-n\sum_{i=1}^{n}y_ix_i-a\sum_{i=1}^{n}x_i^3}{\left(\sum_{i=1}^{n}x_i\right)^2-n\sum_{i=1}^{n}x_i^2}$$

$$c=\frac{\sum_{i=1}^{n}y_i-a\sum_{i=1}^{n}x_i^2-b\sum_{i=1}^{n}x_i}{n}$$

Modelo polinómico general

Utilizando el método de los mínimos cuadrados podemos generalizar, para un conjunto de datos (x_i, y_i) de n elementos, las fórmulas para los parámetros a_j de cualquier función polinómica de grado m de la forma:

$$a_0x^m+a_1x^{m-1}+...+a_jx^{m-j}+...+a_{m-1}x+a_m$$

Elevando al cuadrado las distancias verticales entre la dispersión de puntos y el modelo obtenemos:

$$E=\sum_{n}^{i}\left[y_i-\left(a_0x_i^{\,m}+a_1x_i^{\,m-1}+\ldots+a_jx_i^{\,m-j}+\ldots+a_{m-1}x_i+a_m\right)\right]^2$$

Derivando respecto de cada parámetro a_j se obtiene:

$$\frac{\partial E}{\partial a_0}=-2\sum_{n}^{i}\left[y_i-\left(a_0x_i^{\,m}+a_1x_i^{\,m-1}+\ldots+a_jx_i^{\,m-j}+\ldots+a_{m-1}x_i+a_m\right)\right]x_i^{\,m}$$
$$\frac{\partial E}{\partial a_1}=-2\sum_{n}^{i}\left[y_i-\left(a_0x_i^{\,m}+a_1x_i^{\,m-1}+\ldots+a_jx_i^{\,m-j}+\ldots+a_{m-1}x_i+a_m\right)\right]x_i^{\,m-1}$$
$$\ldots$$
$$\frac{\partial E}{\partial a_j}=-2\sum_{n}^{i}\left[y_i-\left(a_0x_i^{\,m}+a_1x_i^{\,m-1}+\ldots+a_jx_i^{\,m-j}+\ldots+a_{m-1}x_i+a_m\right)\right]x_i^{\,m-j}$$
$$\ldots$$
$$\frac{\partial E}{\partial a_{m-1}}=-2\sum_{n}^{i}\left[y_i-\left(a_0x_i^{\,m}+a_1x_i^{\,m-1}+\ldots+a_jx_i^{\,m-j}+\ldots+a_{m-1}x_i+a_m\right)\right]x_i$$
$$\frac{\partial E}{\partial a_m}=-2\sum_{n}^{i}\left[y_i-\left(a_0x_i^{\,m}+a_1x_i^{\,m-1}+\ldots+a_jx_i^{\,m-j}+\ldots+a_{m-1}x_i+a_m\right)\right]$$

Igualamos las derivadas parciales a cero para optimizar la función de error y obtenemos un sistema de ecuaciones de $(m+1)\times(m+1)$ de la siguiente manera:

$$1)\quad a_0\sum_{n}^{i}x_i^{\,2m}+a_1\sum_{n}^{i}x_i^{\,2m-1}+\ldots+a_j\sum_{n}^{i}x_i^{\,2m-j}+\ldots+a_{m-1}\sum_{n}^{i}x_i^{\,m+1}+a_m\sum_{n}^{i}x_i^{\,m}=\sum_{n}^{i}y_ix_i^{\,m}$$
$$2)\quad a_0\sum_{n}^{i}x_i^{\,2m-1}+a_1\sum_{n}^{i}x_i^{\,2m-2}+\ldots+a_j\sum_{n}^{i}x_i^{\,2m-1-j}+\ldots+a_{m-1}\sum_{n}^{i}x_i^{\,m}+a_m\sum_{n}^{i}x_i^{\,m-1}=\sum_{n}^{i}y_ix_i^{\,m-1}$$
$$\ldots$$
$$j+1)\quad a_0\sum_{n}^{i}x_i^{\,2m-j}+a_1\sum_{n}^{i}x_i^{\,2m-j-1}+\ldots+a_j\sum_{n}^{i}x_i^{\,2m-2j}+\ldots+a_{m-1}\sum_{n}^{i}x_i^{\,m-j+1}+a_m\sum_{n}^{i}x_i^{\,m-j}=\sum_{n}^{i}y_ix_i^{\,m-j}$$
$$\ldots$$
$$m)\quad a_0\sum_{n}^{i}x_i^{\,m+1}+a_1\sum_{n}^{i}x_i^{\,m+2}+\ldots+a_j\sum_{n}^{i}x_i^{\,m-j+1}+\ldots+a_{m-1}\sum_{n}^{i}x_i^{\,2}+a_m\sum_{n}^{i}x_i=\sum_{n}^{i}y_ix_i$$
$$m+1)\quad a_0\sum_{n}^{i}x_i^{\,m}+a_1\sum_{n}^{i}x_i^{\,m}+\ldots+a_j\sum_{n}^{i}x_i^{\,m-j}+\ldots+a_{m-1}\sum_{n}^{i}x_i+a_mn=\sum_{n}^{i}y_i$$

Las soluciones del sistema para el parámetro a_j serían:

$$
a_j = \frac{\begin{array}{c}
\begin{matrix} 1 & 2 & \ldots & j & \ldots & m & m+1 \end{matrix} \\
\begin{vmatrix}
\sum_n^i x_i^{2m} & \sum_n^i x_i^{2m-1} & \ldots & \sum_n^i y_i x_i^{m} & \ldots & \sum_n^i x_i^{m+1} & \sum_n^i x_i^{m} \\
\sum_n^i x_i^{2m-1} & \sum_n^i x_i^{2m-2} & \ldots & \sum_n^i y_i x_i^{m-1} & \ldots & \sum_n^i x_i^{m} & \sum_n^i x_i^{m-1} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\sum_n^i x_i^{2m-j} & \sum_n^i x_i^{2m-j-1} & \ldots & \sum_n^i y_i x_i^{m-j} & \ldots & \sum_n^i x_i^{m-j+1} & \sum_n^i x_i^{m-j} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\sum_n^i x_i^{m+1} & \sum_n^i x_i^{m} & \ldots & \sum_n^i y_i x_i & \ldots & \sum_n^i x_i^{2} & \sum_n^i x_i \\
\sum_n^i x_i^{m} & \sum_n^i x_i^{m-1} & \ldots & \sum_n^i y_i & \ldots & \sum_n^i x_i & n
\end{vmatrix}
\end{array}}{\begin{vmatrix}
\sum_n^i x_i^{2m} & \sum_n^i x_i^{2m-1} & \ldots & \sum_n^i x_i^{2m-j} & \ldots & \sum_n^i x_i^{m+1} & \sum_n^i x_i^{m} \\
\sum_n^i x_i^{2m-1} & \sum_n^i x_i^{2m-2} & \ldots & \sum_n^i x_i^{2m-j-1} & \ldots & \sum_n^i x_i^{m} & \sum_n^i x_i^{m-1} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\sum_n^i x_i^{2m-j} & \sum_n^i x_i^{2m-j-1} & \ldots & \sum_n^i x_i^{2m-2j} & \ldots & \sum_n^i x_i^{m-j+1} & \sum_n^i x_i^{m-j} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\sum_n^i x_i^{m+1} & \sum_n^i x_i^{m} & \ldots & \sum_n^i x_i^{m-j+1} & \ldots & \sum_n^i x_i^{2} & \sum_n^i x_i \\
\sum_n^i x_i^{m} & \sum_n^i x_i^{m-1} & \ldots & \sum_n^i x_i^{m-j} & \ldots & \sum_n^i x_i & n
\end{vmatrix}}
$$

Ajuste de funciones por sustitución

Primer método (Mínimos cuadrados)

Tomando la definición de error medio cuadrático que corresponde a

$$E(a) = \sum_{k]=1}^{n} (f(x) - y_k)^2 ,$$

donde a es el parámetro libre a minimizar, ajustaremos los datos, $(x_1, y_1), \ldots, (x_n, y_n)$ por el método de mínimos cuadrados, para la función

$$f(x) = c \cdot e^{ax} ,$$

el valor que debemos minimizar es:

$$E(a,c) = \sum_{k=1}^{n} (c \cdot e^{ax} - y_k)^2 ,$$

donde a y c son los parámetros que vamos a elegir.

Igualemos a cero las dos derivadas parciales de $E(a,c)$, de esta manera llegamos al sistema de las dos ecuaciones normales, las cuales no son lineales para a y c.

$$c\sum_{k=1}^{n} x_k e^{2ax_k} - \sum_{k=1}^{n} x_k y_k e^{ax_k} = 0$$
$$c\sum_{k=1}^{n} e^{ax_k} - \sum_{k=1}^{n} y_k e^{ax_k} = 0$$

Segundo método (Linealización de los datos)

Este método se utiliza para ajustar diferentes tipos de curvas con un cambio de variable adecuado

Sabiendo que $y_k > 0$ y utilizando logaritmos obtenemos: $\ln y = ax + \ln c$ (1)
Luego, hacemos el siguiente cambio de variable:

$$u = x$$
$$v = \ln y$$
$$b = \ln c$$

Si reemplazamos los anteriores valores en (1) obtenemos:

$$v = au + b$$

De esta manera tenemos los datos transformados, $(u_k, v_k) = (x_k, \ln(y_k))$ para *k=1,…,m*. Donde el problema se puede resolver usando la ecuación de una recta.

Cambios de variables para linealizar datos

A continuación se muestra una tabla en la cual se plantean los cambios que se deben realizar a las variables de acuerdo al tipo de función:

Función $y = f(x)$	Cambios de variables
$y = \frac{a}{x} + b$	$u = \frac{1}{x}, v = y$
$y = \frac{d}{x+c}$	$u = xy, v = y$
$y = \frac{1}{ax+b}$	$u = x, v = \frac{1}{y}$
$y = \frac{x}{ax+b}$	$u = \frac{1}{x}, v = \frac{1}{y}$
$y = a \ln x + b$	$u = \ln x, v = y$
$y = c \cdot e^{ax}$	$u = x, v = \ln y$
$y = \frac{1}{(ax+b)^2}$	$u = x, v = \frac{1}{\sqrt{y}}$
$y = c \cdot x \cdot e^{dx}$	$u = x, v = \ln(\frac{y}{x})$
$y = \frac{l}{1+c \cdot e^{ax}}$	$u = x, v = \ln\left(\frac{l}{y} - 1\right)$

2.2.2.Dimensión cognitiva y didáctica

Esta dimensión se relaciona con la cognición, con la comprensión y el aprendizaje de los estudiantes. Además, se relaciona con la didáctica, concierne con los procesos de acción, reflexión, gestión, diseño y evaluación de la enseñanza de las matemáticas.

2.2.2.1. Reconstrucción enseñanza y aprendizaje

La enseñanza de la estadística genera un aporte importante a la educación matemática, ya que permite "*desarrollar el razonamiento a partir de datos empíricos inciertos. Este tipo de pensamiento estadístico debería ser parte del equipamiento mental de todo ciudadano inteligente*" (PISA, 2003), además promueve el desarrollo de procedimientos, que según el [MEN] Ministerio de Educación Nacional (1998), ayudan a cuantificar, proponer leyes para controlar y elaborar modelos para dar explicación a situaciones, que son consideradas como regidas por el azar, y en consecuencia denominadas aleatorias. Este tipo de acciones se asumen cuando se tiene un tratamiento de situaciones, en donde se dé sentido y se interprete la recolección, organización y representación de datos, permitiendo la toma de decisiones y posteriores predicciones.

El tratamiento de estas situaciones debe estar enmarcado en un contexto determinado y se debe abordar en el aula de clase de diferentes formas, es decir, haciendo uso de recursos y materiales didácticos manipulativos y de simulación. Respecto a este último, el MEN (1998) indica que el uso de los computadores ha hecho más accesible e importante las distintas temáticas de las matemáticas (geometría, probabilidad, estadística y álgebra), ya que "*las nuevas tecnologías amplían el campo de indagación sobre el cual actúan las estructuras cognitivas que se tienen, enriquecen el currículo* *con las nuevas pragmáticas asociadas y lo llevan a evolucionar*" (Ministerio de Educación Nacional, 1998).

En el caso de la estadística, de acuerdo con Godino (1995) la exploración en los computadores no está restringida a la resolución de problemas, sino que se extiende al de los conceptos abstractos, mediante "la creación de micro mundos estocásticos con los que se puede interactuar" (Godino, 1995). Dicha exploración hace posible la experimentación, observación y exploración de diferentes objetos estadísticos, a través de las simulaciones de representaciones gráficas que no son fáciles de captar sin ayuda de los ordenadores.

Esto se evidencia desde hace algunos años donde, según (López y otros, s.f.), el análisis de datos estaba reservado a profesionales, quienes diseñaba sus propios programas con el fin de realizar los cálculos. Sin embargo, estos autores afirman que en los últimos veinte años, se ha visto e impulsado una transformación radical en el campo de la estadística, gracias al desarrollo de programas de computadora diseñados para el análisis estadístico. De la misma manera, indican que "Durante los ochenta, el software estadístico experimentó una vasta revolución tecnológica. Además de las mejoras manifestadas en actualizaciones periódicas, la disponibilidad de computadoras personales condujo al desarrollo de nuevos paquetes que usaban una interfaz manejada por menús" (Berenson & Levine, citado por López y otros, s.f.)

En consecuencia, "el uso de programas de simulación permite poner en manos del alumno un nuevo instrumento que hace posible la exploración y el descubrimiento de conceptos y principios que de otro modo serían mucho más abstractos, contribuyendo a paliar el problema de la falta de experiencia estocástica y a la mejora de la intuición probabilística que, en general, no se desarrollan espontáneamente" (Fishbein, 1975, citado por Godino, 1995)

En este proceso de manipulación de las simulaciones de fenómenos aleatorios existen dos principios que Heitele (1975, citado por Godino, 1995) denomina como: la concepción frecuencia de la probabilidad y el estudio de fenómenos aleatorios. Respecto a este último, se le puede pedir a estudiante que realice muestreos repetidos, representando a través de gráficas las distribuciones obtenidas por medio del ensayo y error, obteniendo en los estudiantes el reconocimiento de que la estimación de una variable cambia de una muestra a otra, ajustándose a un valor predecible.

Es así como los estudiantes pueden, a través de las gráficas y simuladores, según Batanero, Estepa & Godino (1991), estudiar las relaciones de dos variables, ajustar los puntos a una línea recta, estudiar los estadísticos, comparar la línea con los residuos, estudiar la significación estadística del coeficiente de correlación u otros parámetros para descubrir si la relación entre las variables se debe o no al azar; sin embargo, aunque los resultados estadísticos calculados presenten un valor significativo en estadística, la relación entre las variables puede que no se ajuste bien a una línea recta; es aquí donde el estudiante se dará cuenta que el modelo escogido no es el adecuado o el esperado y tendrá que buscar otro modelo que se ajuste a la situación que esté abordando.

Además del uso de simuladores se considera que en el proceso de enseñanza – aprendizaje de la estadística es importante el planteamiento y solución de problemas reales, ya que según Batanero (2002, citado Chaves, 2007), debe existir una relación entre el desarrollo de conceptos estadísticos y la realidad. De acuerdo con esto la autora plantea los dos siguientes puntos, que se deben tener en cuenta en el proceso de construcción de un concepto de dicha área:

“**a).** Que los alumnos lleguen a comprender y a apreciar el papel de la Estadística en la sociedad, incluyendo sus diferentes campos de aplicación y el modo en que la Estadística ha contribuido a su desarrollo. **b).** Que los alumnos lleguen a comprender y a valorar el método estadístico, esto es, la clase de preguntas que un uso inteligente de la Estadística puede responder, las formas básicas de razonamiento estadístico, su potencia y limitaciones.” (Batanero, 2002, citado Chaves, 2007)

Sin embargo, para poder solucionar situaciones relacionadas con la regresión y la correlación es necesario realizar procesos similares en la construcción de conceptos matemáticos, en otras palabras, “la instrucción formal en problemas matemáticos favorece claramente el rendimiento en la solución de problemas de correlación y regresión”. (Perez, M. & Carretero, M., 1990)

De esta manera, la apropiación de conceptos estadísticos, en particular de regresión lineal y no lineal, debe iniciarse desde procesos matemáticos desarrollados con anterioridad.

2.2.3. Dimensión curricular

Se relaciona con el entorno en el que se desenvuelve el proceso de enseñanza-aprendizaje, no se refiere de manera exclusiva al contexto sino también hacen parte de él las políticas educativas bajo las que se rige la educación estadística en nuestro país.

2.2.3.1. Reconstrucción del entorno educativo

Para realizar esta reconstrucción se acude a los textos que hacen parte de las orientaciones curriculares de matemáticas, entre ellos encontramos los lineamientos curriculares y los estándares de calidad, donde se hará particular énfasis en el pensamiento aleatorio y sistemas de datos.

En primer lugar, el Ministerio de Educación Nacional [MEN] (1998), define la probabilidad y la estadística como ramas de las matemáticas que desarrollan procedimientos para cuantificar, las cuales proponen leyes para controlar y elaboran modelos para explicar situaciones que por presentar múltiples variables y de efectos impredecibles, son consideradas como regidas por el azar, y por tanto denominadas aleatorias. El carácter globalizante de la probabilidad y la estadística está en la presencia del pensamiento aleatorio para la comprensión de fenómenos de la vida cotidiana y de las ciencias.

En consecuencia, el MEN (1998), orientó la educación estocástica hacia la resolución de problemas, pues como bien se indica en los lineamientos, la búsqueda de respuestas a preguntas que expliquen fenómenos reales y cotidianos, hace que los estudiantes den sentido a los conceptos estocásticos propios de su nivel.

Es así como se definen procesos y acciones específicas, consideradas como logros, que permiten establecer los productos cognitivos que alcanzan efectivamente los estudiantes, dentro de los cuales se contempla “explorar e interpretar los datos, relacionarlos con otros, conjeturar, buscar configuraciones cualitativas, tendencias, oscilaciones, tipos de crecimiento, buscar correlaciones, distinguir correlación de causalidad, calcular correlaciones y su significación, hacer inferencias cualitativas, diseños, pruebas de hipótesis, reinterpretar los datos, criticarlos, leer entre líneas, hacer simulaciones, saber que hay riesgos en las decisiones basadas en inferencias” (Vasco, 1994, citado por MEN, 1998).

Con el fin de potenciar todas estas acciones y procesos, el MEN (2006) propone unos criterios claros, que son el punto de referencia de lo que el estudiante puede estar en capacidad de *hacer* y *saber hacer,* según el nivel y el tipo de pensamiento. Esto es, la conformación de unos *estándares básicos de competencias en matemáticas* (MEN, 2006), los cuales a partir de cinco campos de pensamientos (el numérico, el espacial, el métrico o de medida, el variacional y el aleatorio o probabilístico) pretenden abordar todas la competencias, capacidades y aptitudes en el área de matemáticas. Cada uno de los cinco pensamientos son profundizados mediante niveles de curso que están determinados en grupos así: 1° a 3°, 4° a 5°, 6° a 7°, 8° a 9° y 10° a 11°.

Es así como para el grado 11° se establecen estándares relacionados con los cinco tipos de pensamiento. Para el presente trabajo se hará especial énfasis en el pensamiento *Aleatorio y Sistema de Datos,* también llamado *probabilístico o estocástico,* que de acuerdo con el MEN (2006) ayuda a tomar decisiones en situaciones de incertidumbre, de azar, en situaciones en las que no se puede predecir qué va a ocurrir. Tiene su fundamento en conceptos y procedimientos de la teoría de probabilidades y de la estadística inferencial, e indirectamente en la estadística descriptiva y en la combinatoria. Permite, a través de la exploración y la investigación, solucionar problemas y modelar situaciones físicas, sociales tan probabilísticas como estadísticas.

Es así como en el pensamiento aleatorio y sistema de datos se pretende que el estudiante domine "los conceptos y procedimientos necesarios para recoger, estudiar, resumir y diagramar sistemas de datos estadísticos y tratar de extraer de ellos toda la información posible con la ayuda de calculadoras, hojas de cálculo y otros programas de análisis de datos, con el fin de intentar predecir dentro de ciertos rangos el curso de los acontecimientos respectivos y de tomar decisiones lo más razonables posibles ante la imposibilidad de saber con certeza lo que va a pasar"

Basados en lo anterior, a continuación se dan a conocer los estándares, de grado once del pensamiento aleatorio y sistema de datos, bajo los cuales se rige la presente propuesta:

- Interpreto y comparo resultados de estudios con información estadística proveniente de medios de comunicación.

- Justifico o refuto inferencias basadas en razonamientos estadísticos a partir de resultados de estudios publicados en los medios o diseñados en el ámbito escolar.

- Diseño experimentos aleatorios (de las ciencias físicas, naturales o sociales) para estudiar un problema o pregunta.

- Describo tendencias que se observan en conjuntos de variables relacionadas.

- Uso comprensivamente algunas medidas de centralización, localización, dispersión y correlación (percentiles, cuartiles, centralidad, distancia, rango, varianza, covarianza y normalidad).

Teniendo en cuenta estos criterios que establece el MEN (2006), se plantean las diferentes situaciones que estarán plasmadas en los aplicativos que permitirán explorar los diferentes modelos de regresión lineal y no lineal.

CAPÍTULO 3. METODOLOGÍA Y DISEÑO DE APLICATIVOS Y SIMULADORES

3.1. INTRODUCCIÓN

Teniendo en cuenta que el desarrollo de la propuesta debe estar sustentada bajo unos referentes metodológicos, en el presente capítulo se abordan los contenidos relacionados con las diversas formas en las cuales se fue construyendo el presente trabajo, es decir, se da a conocer la metodología utilizada, la cual no corresponde a una metodología de investigación, sino a un proceso de indagación referido al diseño de aplicativos y simuladores de modelos de regresión lineal y no lineal.
De la misma manera, en una segunda parte, se explica detalladamente el diseño de los simuladores y aplicativos, donde además se da a conocer el desarrollo y la implicación didáctica de los mismos en el aprendizaje y enseñanza de la estadística y en particular del objeto estadístico en estudio.

3.2. METODOLOGÍA

La metodología utilizada en la presente propuesta no está sustentada bajo un referente teórico determinado, ya que corresponde un proceso de indagación y no a uno de investigación. Bajo estas condiciones, se establece unos elementos metodológicos, que están ligados a la forma como se construye el trabajo y los tiempos en los cuales se ejecuta cada uno de ellos.

3.2.1. Descripción metodológica

Para iniciar, se describe el proceso primario de indagación, donde se realizó la búsqueda de antecedentes, que fueron organizados en tres grupos o dimensiones, denominadas *epistémica, enseñanza – aprendizaje* y *curricular*. La primera hace referencia a los aspectos netamente del objeto matemático, su definición y su historia; la segunda se refiere a los procesos y dificultades referidos al profesor y a los estudiantes; y el tercero se refiere a las leyes y orientaciones curriculares que se han propuesto en nuestro país en el área de matemáticas.

Mientras se realizaba la búsqueda de antecedentes se procedía a elaborar formular el interés de estudio, las preguntas (orientadora y auxiliares) y los objetivos (general y específicos) de la propuesta. Gracias a la elaboración y búsqueda de antecedentes se pudo delimitar el tema de interés, demarcando un poco más la finalidad de realizar el trabajo.

Ya delimitado el tema y teniendo unos objetivos trazados, se inició con la construcción del marco teórico – conceptual, el cual se organizó a partir de la dimensiones planteadas en los antecedentes. Es así como en el marco conceptual se realizó una reconstrucción de la regresión lineal y no lineal, una reseña histórica de este objeto estadístico, una descripción de las orientaciones curriculares de nuestro país y la forma como éstas proponen el trabajo del objeto de estudio en las aulas de matemáticas, y los procesos desarrollados tanto por el profesor como por el estudiante al momento de abordar los modelos de regresión lineal y no lineal.

A medida que se iba elaborando el marco teórico, se estaban construyendo los primeros aplicativos en el programa Descartes. A su vez, se iban planteando las situación

es que posibilitan la exploración de los modelos de regresión lineal y no lineal, a través de dichos aplicativos.

Teniendo el marco teórico y algunos aplicativos construidos, se centra la atención en el diseño de los aplicativos y simuladores, donde inicialmente se evalúan las situaciones creadas, a partir de los estándares curriculares para el grado 11°, para luego continuar con el diseño de los simuladores, al estilo de una página Web. Durante dicho diseño se tuvo en cuenta la esquematización utilizada, la facilidad y entendimiento del manejo de software y la pertinencia en la exploración de los modelos de regresión lineal y no lineal.

Hecho este proceso y bajo algunos cambios y sugerencias recibidas, se culmina el trabajo referente al diseño de los aplicativos y simuladores, para luego reflexionar y analizar acerca de la importancia y pertinencia del uso de los simuladores para ilustrar y explorar los modelos de regresión lineal y no lineal.

3.3. DISEÑO DE SIMULADORES Y APLICATIVOS

3.3.1. Caracterización de los Simuladores y aplicativos de los modelos de regresión lineal y no lineal

Los simuladores y aplicativos de los modelos de regresión lineal y no lineal, son un software educativo creado con ayuda del programa Descartes, el cual permite simular problemas que requieran el manejo de los modelos de regresión lineal, mínimos cuadrados, cuadrático, potencial, hiperbólico, exponencial y logarítmico.

Dichos simuladores y aplicativos se desarrollan con varios fines, que están implícitos dentro del manejo del software, pero que por descripción del mismo se darán a conocer a continuación.

3.3.2. Objetivos de los simuladores y aplicativos

Los objetivos de los simuladores y aplicativos se construyen a partir de la organización de las dimensiones (*epistémica, enseñanza – aprendizaje* y *curricular)* propuestas para la búsqueda de antecedentes y construcción del marco teórico – conceptual en los capítulos 1 y 2, respectivamente. Es así, como se plantea un objetivo general que engloba las tres dimensiones mencionadas y cuatro objetivos específicos, los dos primeros enfocados a la dimensión enseñanza aprendizaje, el siguiente a la epistémica y el último a la dimensión curricular.

3.3.2.1. Objetivo General

Proporcionar herramientas, al profesor y al estudiante, que permitan ilustrar y comparar los diferentes modelos de regresión lineal y no lineal, para que sean explorados y apropiados por estudiantes de grado 11°, a través de la solución de situaciones.

3.3.2.2. Objetivos específicos

- Proporcionar al estudiante las herramientas necesarias para elegir un modelo de regresión que le permita dar solución a una situación

- Proporcionar al profesor una herramienta educativa que le permita orientar y acompañar los procesos relacionados con la exploración y uso de los modelos de regresión en la solución de situaciones.

- Exponer la reconstrucción histórica y teórica del objeto estadístico para que el estudiante se apropie y conozca la evolución del mismo

- Abordar las competencias curriculares, propuestas por el MEN (2006), relacionadas con la regresión lineal y no lineal

3.3.3. Diagramación de los simuladores y aplicativos

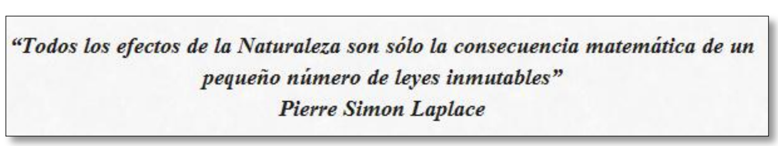
El software de simuladores y aplicativos de los modelos de regresión lineal y no lineal, tiene una estructura basada en la organización de imágenes, de representaciones gráficas y simuladores, de reconstrucción teórica e histórica y de aplicaciones a situaciones. A continuación se explica la forma en que está diseñado cada uno de estos componentes y su posible influencia a la hora de abordar los modelos de regresión lineal y no lineal.

3.3.3.1. Presentación del sitio web

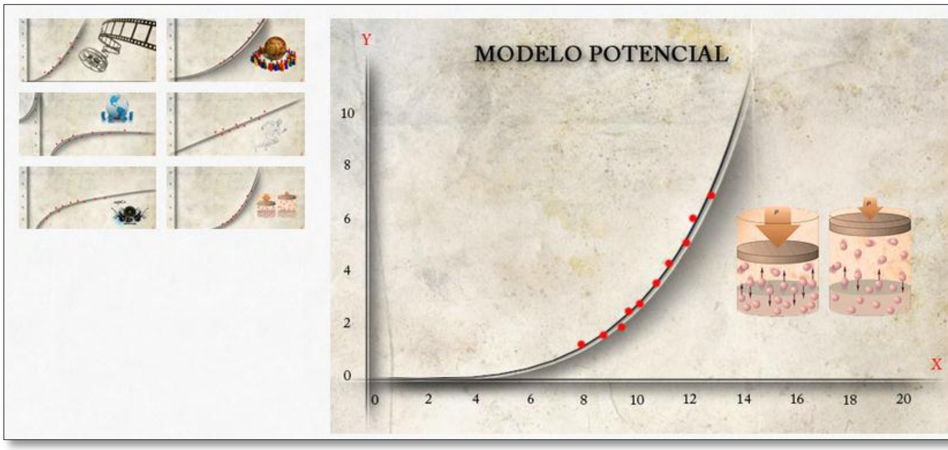
La siguiente corresponde a una descripción de la diagramación que compone el software:

- **Barra botones de acción:** Al inicio del software, en la parte superior, se presenta una barra de botones que están distribuidos de la siguiente manera: reconstrucción histórica, los modelos de regresión (mínimos cuadrados, lineal, cuadrático, potencial, hiperbólico, exponencial y logarítmico) y las aplicaciones de dichos modelos en situaciones.

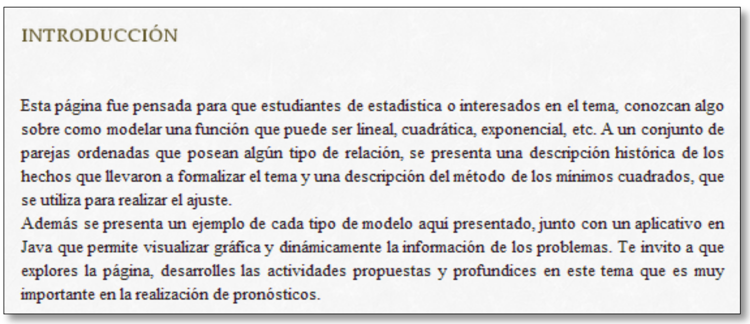
- **Frase matemática**: ***"****Todos los efectos de la Naturaleza son sólo la consecuencia matemática de un pequeño número de leyes inmutables"* de *Pierre Simon Laplace*, la cual pretende hacer ver al estudiante la influencia e importancia de las matemáticas en la interacción con su mundo real, teniendo en cuenta que uno de los objetivos del software es la aplicación de los modelos de regresión en situaciones.



- **Gráficas de los modelos:** muestran las gráficas de los modelos lineal, cuadrático, potencial, hiperbólico, exponencial y logarítmico, cada una acompañada de una imagen relacionada con el tipo de función que se está trabajando en cada caso.



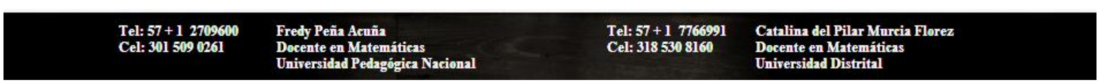
- **Introducción**: Es un espacio donde se da la bienvenida al lector, explicando la finalidad de los aplicativos y simuladores; se muestra de manera general y reducida una descripción del sitio y sus distintas opciones de uso.



INTRODUCCIÓN

Esta página fue pensada para que estudiantes de estadística o interesados en el tema, conozcan algo sobre como modelar una función que puede ser lineal, cuadrática, exponencial, etc. A un conjunto de parejas ordenadas que posean algún tipo de relación, se presenta una descripción histórica de los hechos que llevaron a formalizar el tema y una descripción del método de los mínimos cuadrados, que se utiliza para realizar el ajuste.
Además se presenta un ejemplo de cada tipo de modelo aquí presentado, junto con un aplicativo en Java que permite visualizar gráfica y dinámicamente la información de los problemas. Te invito a que explores la página, desarrolles las actividades propuestas y profundices en este tema que es muy importante en la realización de pronósticos.

Nombres de los autores: sus profesiones y sus respectivos números de teléfono, para que los lectores puedan contactarlos. Estos son mostrados en todas las páginas y pantallas del software.

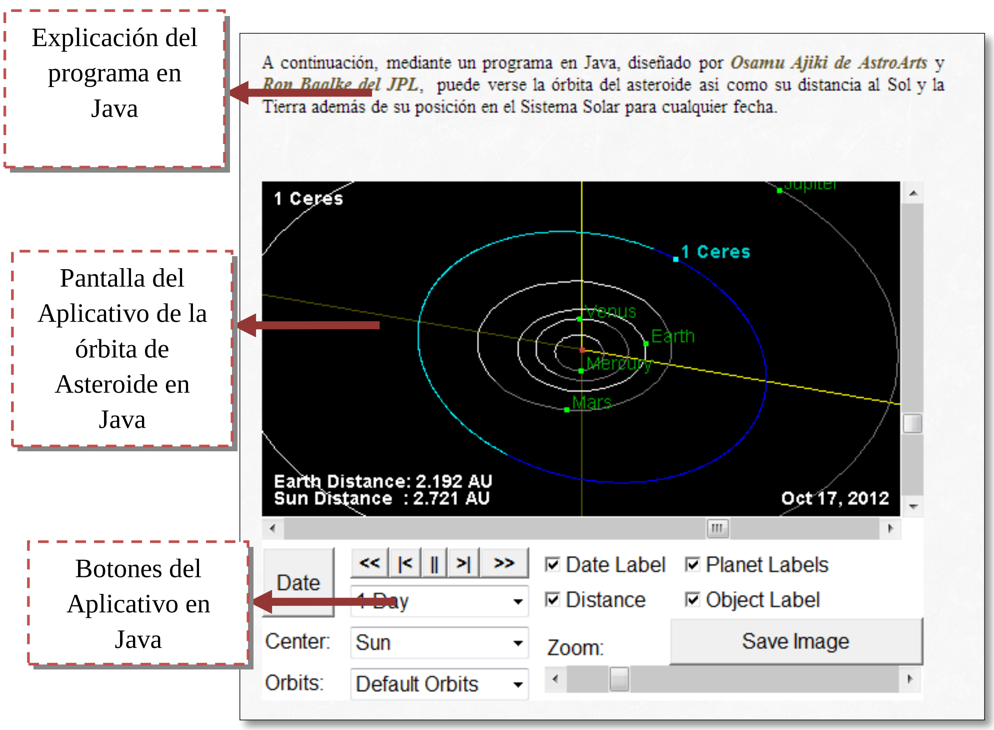


3.3.3.2. Exploración de los botones de acción

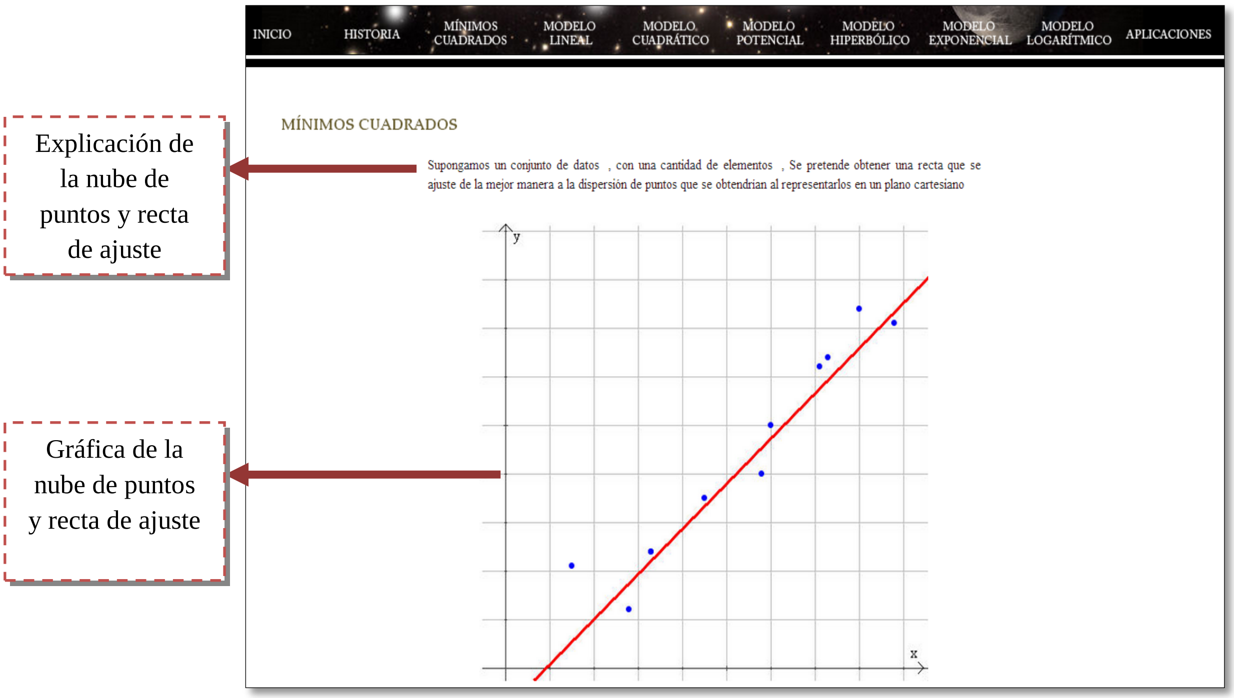
A continuación describe algunos de los botones de la *barra de botones de acción*, con el fin de explicar al lector el contenido que presenta cada botón.

- **Botón Historia**. En la primera parte del botón de historia el estudiante encuentra una reseña acerca de la construcción y desarrollos que llevaron a cabo Piazzi y Gauss, frente al tratamiento de temáticas de regresión, en particular de regresión lineal. También puede observar las fotos de los dos personajes nombrados.

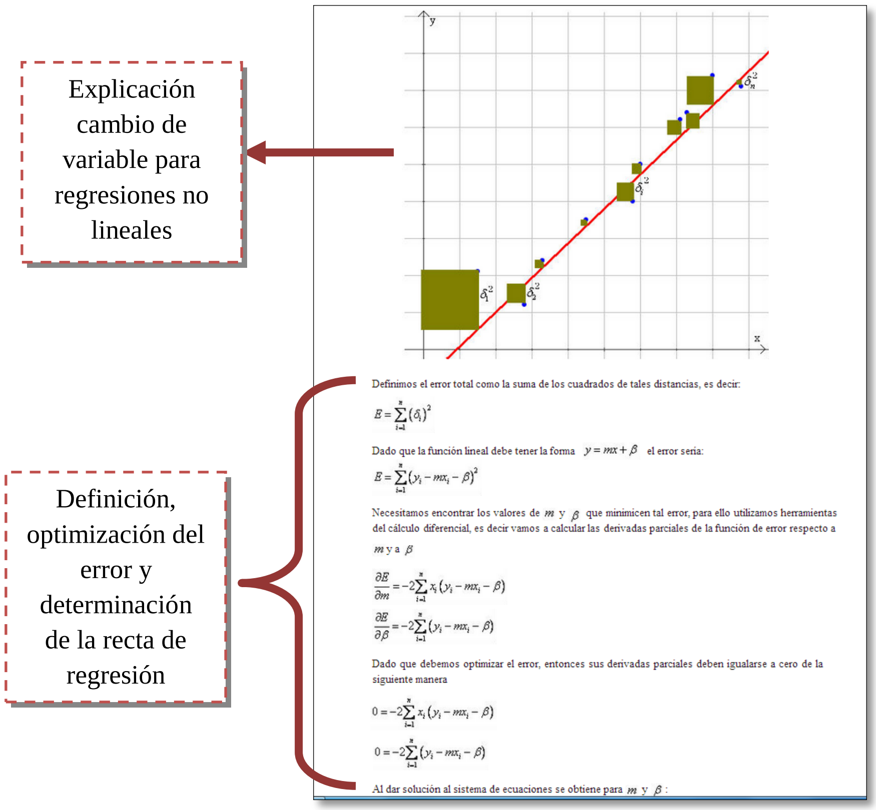
Posteriormente, el estudiante encuentra una segunda parte del botón de *historia*, donde puede explorar un programa en Java, el cual le permite ver la órbita del asteroide, determinando su distancia al sol y a la tierra y su posición en el sistema solar para cualquier fecha.



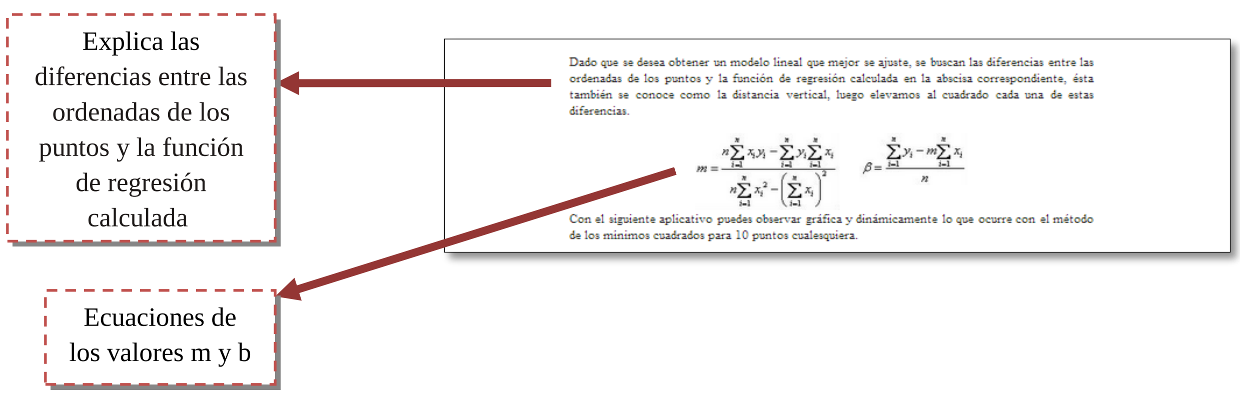
- **Botón Mínimos cuadrados**. En este botón el estudiante encuentra una explicación del propósito de utilizar el método de mínimos cuadrados en el proceso de linealización. Para lo cual, puede observar una gráfica de una nube de puntos con una recta en el plano, con la que dotará de sentido el uso de dicho método.

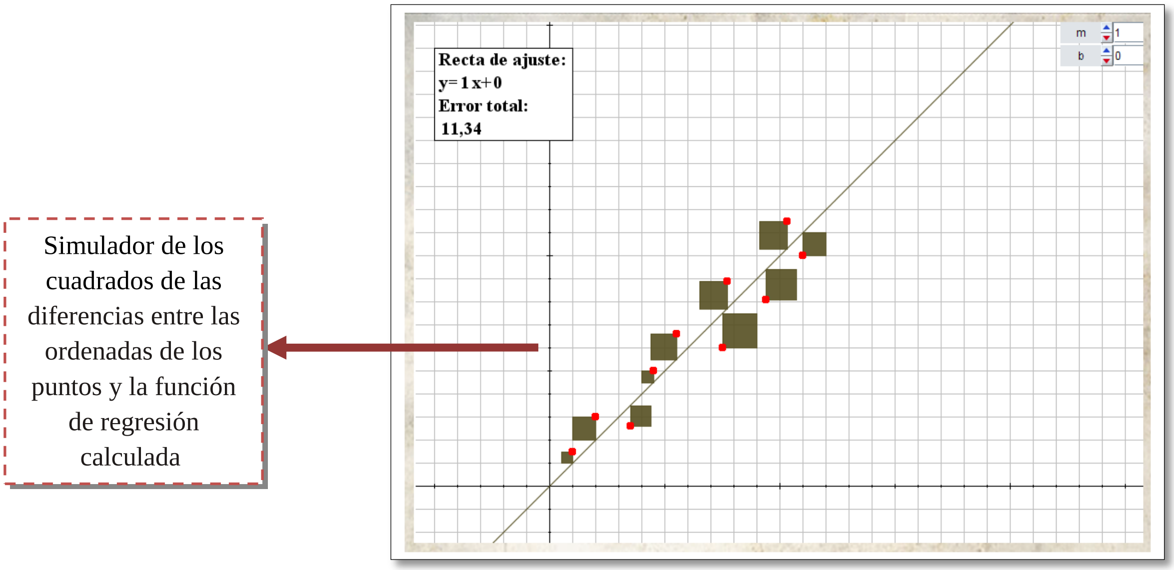


Posteriormente se muestra, a través de una gráfica y de ecuaciones, con el fin de determinar el error total y los valores de m y b.



Luego, el estudiante tiene la opción de observar las ecuaciones utilizadas para determinar los parámetros *m* y *b* en funciones lineales y de manipular una gráfica a través de la cual puede explorar el método para 10 puntos cualesquiera.





Por último, el estudiante encontrará una tabla que indica el cambio de variable que se debe realizar para modelos no lineales.

Explicación cambio de variable para regresiones no lineales

Para regresiones no lineales basta realizar un cambio de variables a fin de linealizar las ecuaciones resultantes, dado que los sistemas de ecuaciones no lineales que se obtienen, no son fáciles de resolver. A continuación algunos cambios de variables que se usan para modelos no lineales.

Tabla de cambio de variable para regresiones no lineales

Función $y = f(x)$	Cambios de variables
$y = \frac{a}{x} + b$	$u = \frac{1}{x}, v = y$
$y = \frac{d}{x+c}$	$u = xy, v = y$
$y = \frac{1}{ax+b}$	$u = x, v = \frac{1}{y}$
$y = \frac{x}{ax+b}$	$u = \frac{1}{x}, v = \frac{1}{y}$
$y = a \ln x + b$	$u = \ln x, v = y$
$y = c \cdot e^{ax}$	$u = x, v = \ln y$
$y = \frac{1}{(ax+b)^2}$	$u = x, v = \frac{1}{\sqrt{y}}$
$y = c \cdot x \cdot e^{dx}$	$u = x, v = \ln(\frac{y}{x})$
$y = \frac{l}{1+c \cdot e^{ax}}$	$u = x, v = \ln\left(\frac{l}{y} - 1\right)$

- **Botón Modelos de regresión.** En los diferentes botones de modelos de regresión, el estudiante puede explorar cada una de las representaciones gráficas de dichos modelos. Teniendo en cuenta que el diseño de la información contenida en los botones de los modelos de regresión (lineal, cuadrático, potencial, hiperbólico, exponencial y logarítmico) es el mismo, se explicará a continuación la estructura de uno de ellos, para que sea asimilada y entendida para los demás botones.

Modelo de regresión lineal

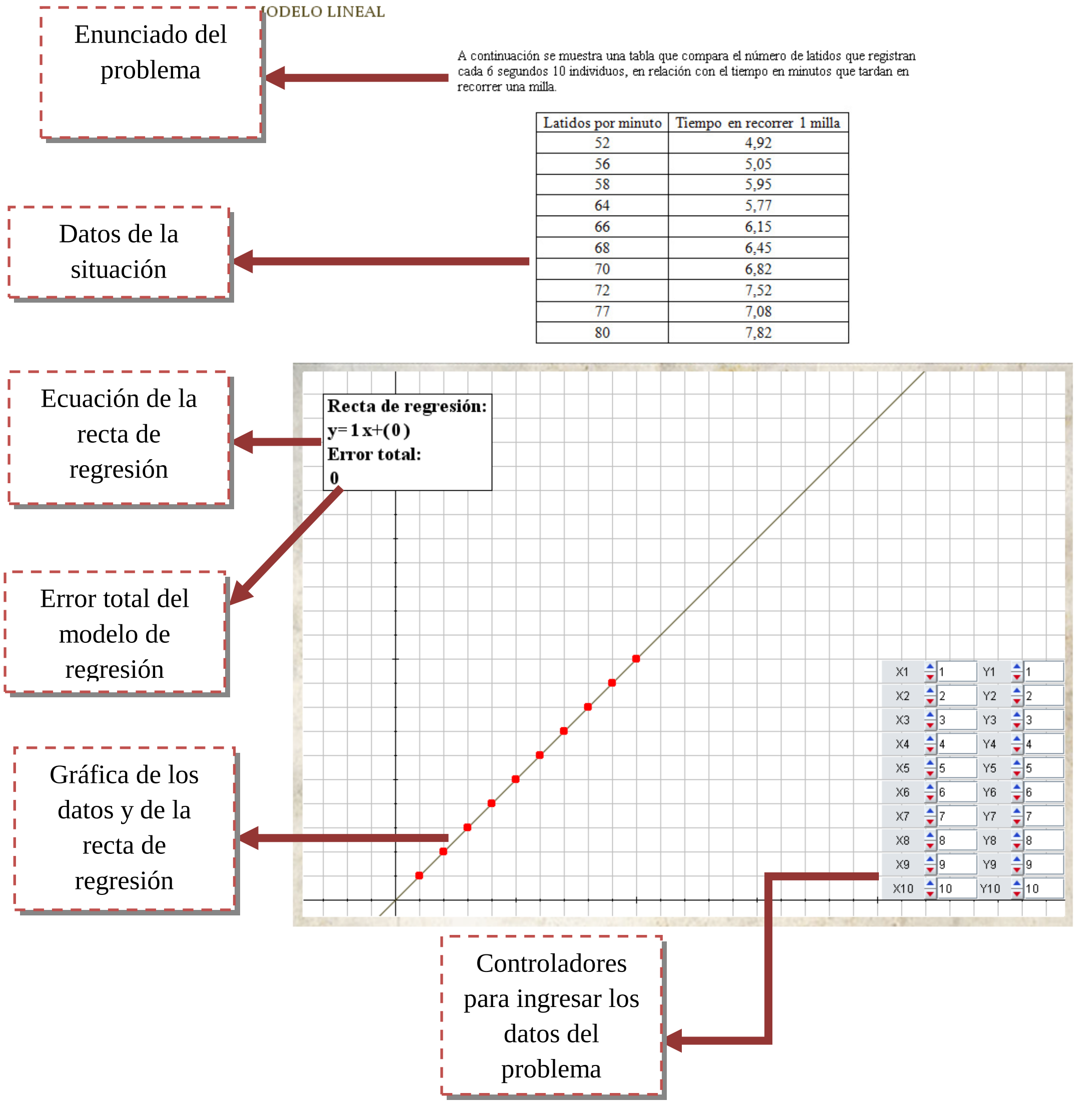
Se le presenta al estudiante una actividad, a la cual le dará solución y explorará a través de la interacción con una gráfica diseñada como simulador de dicha situación. En la gráfica se encuentran algunos elementos que él puede utilizar para hacer pronósticos y concluir frente a la situación propuesta.

El primer elemento muestra la *ecuación de la función de regresión*, que en este caso corresponde a una recta. A través de ésta, el estudiante puede observar la ecuación de la función que está graficando para luego utilizarla en algoritmos o pronósticos que deba desarrollar.

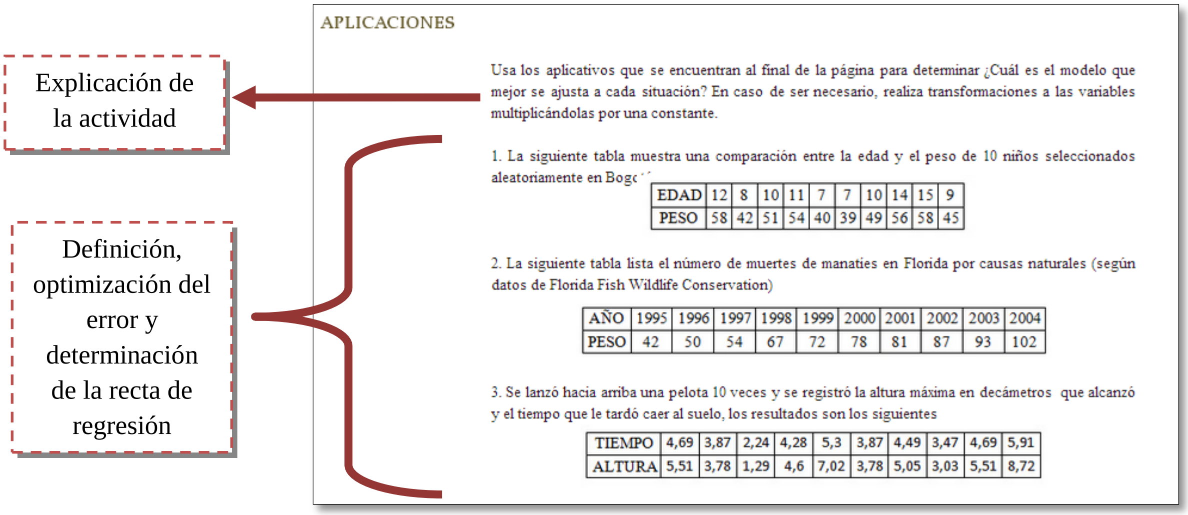
El segundo elemento es el *error total*, por medio del cual el estudiante evita realizar procedimientos analíticos respecto a la suma de los cuadrados de las diferencias entre los valores de la nube de puntos y la función de regresión, dicho valor le indica en cuál de las manipulaciones realizadas, se presenta menor error total.

Los elementos tres y cuatro corresponden, respectivamente, al aplicativo gráfico de la función de regresión y a los botones o controladores de los datos de la situación, los cuales le permiten al estudiante simular la gráfica de acuerdo a los datos que el desee ingresar.

En general, los cuatro elementos permiten solucionar la actividad planteada y reconocer distintas representaciones de los modelos de regresión.



Botón aplicaciones. En este botón el estudiante encuentra tres situaciones, que debe solucionar y explorar haciendo uso de seis aplicativos gráficos, cada uno de los cuales corresponden a los seis modelos de regresión que se manejan en el software.



APLICACIONES

Usa los aplicativos que se encuentran al final de la página para determinar ¿Cuál es el modelo que mejor se ajusta a cada situación? En caso de ser necesario, realiza transformaciones a las variables multiplicándolas por una constante.

1. La siguiente tabla muestra una comparación entre la edad y el peso de 10 niños seleccionados aleatoriamente en Bogc

EDAD	12	8	10	11	7	7	10	14	15	9
PESO	58	42	51	54	40	39	49	56	58	45

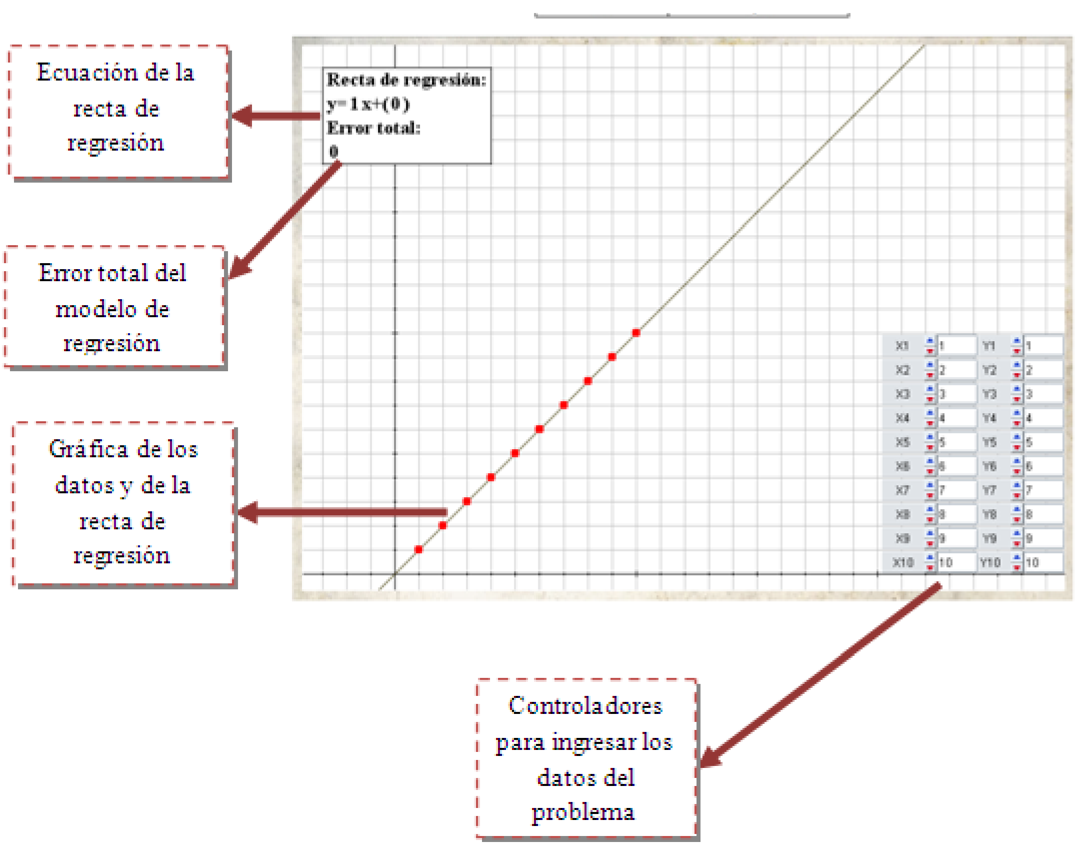
2. La siguiente tabla lista el número de muertes de manaties en Florida por causas naturales (según datos de Florida Fish Wildlife Conservation)

AÑO	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
PESO	42	50	54	67	72	78	81	87	93	102

3. Se lanzó hacia arriba una pelota 10 veces y se registró la altura máxima en decámetros que alcanzó y el tiempo que le tardó caer al suelo, los resultados son los siguientes

TIEMPO	4,69	3,87	2,24	4,28	5,3	3,87	4,49	3,47	4,69	5,91
ALTURA	5,51	3,78	1,29	4,6	7,02	3,78	5,05	3,03	5,51	8,72

La siguiente corresponde a una de las gráficas de los simuladores que el estudiante puede utilizar para solucionar y explorar las situaciones propuestas. Cada gráfica tiene la misma estructura y funcionalidad de los aplicativos mostrados en los botones de los diferentes modelos de regresión.



3.3.4. Guía de trabajo para el estudiante

A continuación se muestra un taller guía para el estudiante, basado en el planteamiento de situaciones reales y cotidianas, a través del cual él podrá explorar los distintos modelos de regresión lineal y no lineal, haciendo uso de los aplicativos y simuladores de dichos modelos.

Esta guía se propone como un elemento orientador y un recurso para el profesor y el estudiante en el aula de matemáticas, el cual debe ser orientado por el docente de tal manera que el alumno vea la necesidad y tenga el interés de utilizar el software para dar solución a las situaciones planteadas y realizar pronósticos de las mismas.

Por otro lado, se proponen una serie de actividades de los modelos de regresión (Ver Anexo 1) las cuales pueden ser trabajadas a través del uso de los aplicativos y simuladores.

GUÍA PARA EL ESTUDIANTE

El géiser Old Faitfhful es la atracción más visitada en el Parque Nacional Yellowstone. Está ubicado cerca del hotel Old Faithful Inn, que tal vez sea la segunda atracción más visitada de Yellowstone. Los turistas disfrutan de la comida, las bebidas, el alojamiento y las tiendas del hotel, pero quieren asegurarse de ver al menos una erupción del famoso géiser Old Faithful. Los guardabosques del parque ayudan a los turistas publicando el momento de la siguiente erupción. ¿Cómo hacen esas predicciones?

1 El géiser Old Faithful del Parque Nacional Yellowstone arroja agua directamente hacia arriba con una altura de hasta 150 pies. © iStockphoto / Zuk

Cuando el Old Faitfhful hace erupción, se registran las siguientes mediciones: duración (en segundos) de la erupción, el intervalo de tiempo (en minutos) entre la erupción anterior y la actual (intervalo previo), el intervalo de tiempo (en minutos) entre la erupción actual y la siguiente (intervalo posterior), y la altura (en pies) de la erupción. En la siguiente tabla se muestran los registros de diez de las erupciones (no consecutivas) del Old Faitfhful.

Erupciones del géiser Old Faitfhful										
Duración	240	120	178	234	235	269	255	220	213	241
Intervalo previo	98	90	92	98	93	105	81	108	94	92
Intervalo posterior	86	65	72	84	83	94	83	80	80	85
Altura	140	110	125	120	140	120	125	150	120	110

Dado que el interés de los trabajadores es predecir el momento en el cual el Old Faitfhful hará erupción

- Realiza las gráficas de intervalo posterior contra duración, intervalo previo y altura, y observa el comportamiento que tiene cada una de estas.

Gráficamente

- ¿Encuentras alguna relación entre el intervalo posterior y el intervalo previo? Justifica

- ¿Encuentras alguna relación entre el intervalo posterior y la altura? Justifica

- ¿Encuentras alguna relación entre el intervalo posterior y la duración? Justifica

- ¿Cuál consideras es el parámetro principal que tienen en cuenta los trabajadores del parque para predecir la siguiente erupción? Justifica

El coeficiente de correlación lineal determina el grado de dependencia lineal de una variable sobre la otra, se define como: $r = \frac{\delta_{XY}}{\delta_X \delta_Y}$ donde δ_{XY} representa la covarianza y δ la varianza, cuanto más cercano sea a 1 o -1, mayor dependencia lineal de una variable sobre la otra existe y cuanto más cercano sea a cero menor dependencia lineal existe.

- Calcula los coeficientes de correlación lineal para el intervalo posterior y los demás parámetros, compara los resultados con lo que observas en las gráficas que construiste ¿Son coherentes con lo que interpretaste gráficamente?

Para obtener una recta $y = mx + \beta$ que se ajuste a la dispersión de datos, se usan las fórmulas:

$$\boxed{m = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}}$$

$$\boxed{\beta = \frac{\sum_{i=1}^{n} y_i - m\sum_{i=1}^{n} x_i}{n}}$$

Siendo los x_i valores para la duración de la erupción y los y_i valores del intervalo posterior.

Duración x_i	**Intervalo posterior** y_i	x_i^2	$x_i y_i$
240	86		
120	65		
178	72		
234	84		
235	83		
269	94		
255	83		
220	80		
213	80		
241	85		
$\sum_{i=1}^{n} x_i =$	$\sum_{i=1}^{n} y_i =$	$\sum_{i=1}^{n} x_i^2 =$	$\sum_{i=1}^{n} x_i y_i =$

- Usa los resultados de la tabla para calcular los parámetros m y β y construir la ecuación de la recta de ajuste.

- Usa el aplicativo de regresión lineal para construir la recta de regresión y comprobar que los parámetros m y β coinciden con lo que calculaste (ten presente realizar cambios de variable para ajustar los datos al aplicativo).

- Predice el tiempo que tardará en volver a hacer erupción el géiser Old Faitfhful si la duración de la erupción anterior fue 100 segundos, repite el proceso para 50, 75, 125, 150, 175 y 200 segundos.

3.3.5. Los estándares de matemáticas en los aplicativos y simuladores

Para la construcción de los aplicativos y simuladores de regresión lineal y no lineal, se tuvo en cuenta los Estándares Básicos de Competencias en Matemáticas de grado 11º, los cuales fueron abordados en la Dimensión curricular del marco teórico - conceptual de la presente propuesta. Cada de uno de los estándares allí expuestos tienen gran importancia y participación dentro de los conceptos, diagramación y en general en el diseño del software.

Estándar	Relación con los aplicativos y simuladores
Interpreto y comparo resultados de estudios con información estadística provenientes de medios de comunicación.	Se utiliza en las situaciones propuestas, ya que a partir de las gráficas de los modelos de regresión, el estudiante debe comparar la información de los problemas planteados y explorar la situación para tomar decisiones o resolver dichas situaciones.
Justifico o refuto inferencias basadas en razonamientos estadísticos a partir de resultados de estudios publicados en los medios o diseñados en el ámbito escolar.	Se evidencia en los problemas planteados y en el uso de los simuladores y aplicativos, ya que el estudiante debe justificar y sacar conclusiones frente a la exploraciones hechas en las gráficas, para dar solución a las situaciones allí propuestas.
Diseño experimentos aleatorios (de las ciencias físicas, naturales o sociales) para estudiar un problema o pregunta.	A pesar de que el estudiante no debe diseñar los experimentos, si debe enfrentarse a situaciones de las ciencias físicas, naturales o sociales, con el fin de encontrar algunas regularidades, determinar el modelo de regresión que se ajusta y realizar pronósticos frente a los problemas propuestos.
Describo tendencias que se observan en conjuntos de variables relacionadas.	Este estándar se evidencia cuando el estudiante debe hacer uso de las gráficas, las ecuaciones de las funciones y el error total para determinar la relación entre las variables de la situación planteada y poder realizar pronósticos y determinar el modelo que mejor se ajusta a la situación planteada.
Uso comprensivamente algunas medidas de centralización, localización, dispersión y correlación (percentiles, cuartiles, centralidad, distancia, rango, varianza, covarianza y normalidad).	Para determinar el modelo que mejor se ajusta a la situación y en particular al conjunto de datos, el estudiante debe comprender y darle significado a conceptos como correlación, regresión y covarianza.

La anterior configuración entre estándares y los aplicativos y simuladores es la que permite hacer un análisis y reflexión frente a la pertenencia e importancia del aplicativo para la consecución y abordaje de los estándares curriculares para el área de matemáticas de grado 11º, relacionados con la regresión lineal y lineal.

CAPÍTULO 4. REFLEXIONES Y CONCLUSIONES

4.1. INTRODUCCIÓN

El proceso de reflexión se lleva a cabo a partir de las tres dimensiones (*enseñanza – aprendizaje, epistémica y curricular)* que han sido abordadas durante la propuesta; de esta manera, se logra enfatizar en tres puntos principales relacionados con los objetivos del trabajo; el primero se refiere al planteamiento de situaciones cotidianas, donde se centra la atención en lo que se espera que el estudiante realice frente a dichas situaciones, el segundo es el análisis del diseño de los simuladores y aplicativos, donde se reflexiona acerca de los aportes que brinda el software en el proceso enseñanza – aprendizaje, y el tercero, analiza y reflexiona respecto a la importancia y pertinencia de los simuladores y aplicativos en el currículo de matemáticas, tanto en los planes de estudio como en las políticas nacionales.

Por último, se concluye frente al uso de las TICS y de los materiales educativos multimedia en la construcción y comprensión de conceptos estadísticos desarrollados en el grado 11º de la educación media, en particular de la Regresión Lineal y no Lineal, y se reflexiona en torno a los aportes del presente trabajo tanto a nivel personal como profesional.

4.2. Reflexión en torno a las situaciones planteadas

El planteamiento de situaciones en los aplicativos y simuladores surge de la necesidad de enfrentar al estudiante a la búsqueda de estrategias y solución actividades, las cuales son fundamentales para la comprensión de conceptos estadísticos, en particular de regresión lineal y no lineal.

Además, se presenta como un instrumento didáctico de ayuda para el profesor, por medio del cual, involucra al estudiante en un ambiente distinto de aprendizaje, que fue propuesto teniendo en cuenta un contexto y una aplicabilidad real.

4.3. Reflexión en torno al diseño de simulaciones y aplicativos

Las simulaciones y aplicativos son propuestos como una herramienta de ayuda para los estudiantes, la cual facilita el análisis de la situación planteada a través de distintos tipos de representación. El primero de ellos corresponde a la ecuación de la función involucrada, permitiendo que el estudiante reconozca la representación analítica de la regresión de cada situación; el segundo es la representación gráfica de la función, que no sólo se presenta para relacionar la representación analítica con ésta, sino para que el estudiante determine el modelo que mejor se ajusta al conjunto de datos y por tanto a la situación.

Por otra parte el software evita que los estudiantes realicen y desarrollen algoritmos extensos y procesos repetitivos, con los cuales posiblemente no logren construir el concepto, ni le den sentido a la situación y mucho menos a la solución de la misma.

4.4. Reflexión en torno a la importancia y pertinencia del diseño de simulaciones y aplicativos

El diseño y construcción de simulaciones y aplicativos acompañados de actividades que pretende ofrecer a la comunidad de educadores matemáticos, y en particular estadísticos, herramientas y materiales educativos que apoyen el desarrollo y ejecución de los planes de estudio propuestos en cada institución educativa y de los estándares matemáticos y políticas nacionales, en la educación media.

Respecto a la ejecución de los planes de estudio, los aplicativos permiten orientar los procesos de enseñanza – aprendizaje, los cuales facilitan el trabajo del profesor y del estudiante y se constituyen como una alternativa para abordar temáticas y competencias que involucren conceptos abstractos, los cuales son de difícil entendimiento sin ayudas didácticas y gráficas oportunas. Posiblemente el aplicativo permite al estudiante darle sentido y significado, es decir, apropiarse del concepto de regresión lineal y no lineal, por medio de la teoría, la historia y algunos ejemplos y aplicaciones de dicho concepto. De esta manera, el anterior proceso, llevado a cabo en el aula de matemáticas, se verá aprovechado en el momento en que el estudiante utilice satisfactoriamente los simuladores, es decir, que los aproveche para determinar comportamientos y hacer pronósticos de las situaciones planteadas.

Por otro lado, en la ejecución y desarrollo de los Estándares Básicos de Competencias en Matemáticas (MEN, 2006) el aplicativo pretende que el estudiante interprete y compare resultados que se obtienen de situaciones del contexto, justificando o dando un contra ejemplo a través del razonamiento estadístico, obtenido de la lectura y el análisis de gráficas y de diversas representaciones de las funciones y ecuaciones que intervienen en los modelos de regresión lineal y no lineal. Es así, como finalmente, de manera implícita o explícita, el estudiante podrá describir relaciones y correspondencias entre conjuntos de variables, aplicados a la estadística, en particular al objeto de estudio.

4.5. Reflexión personal

Empezar un proceso de indagación implica que se cuente con un interés por parte de una persona o un grupo de personas, el cual surge en nuestro caso en la práctica docente y en los diferentes ejes de formación universitaria que involucraban la regresión lineal y no lineal, nuestro interés se centraba en la regresión como tal, pero después de algunas sugerencias de expertos nos dimos cuenta que la regresión además de ser un tema que surge en el grado once de la educación media en Colombia, y que en ocasiones, por experiencias propias, vimos que se ve relegado en los planes de estudio, decidimos, con ayuda del asesor, pensar en una herramienta que pudiera mediar los procesos de enseñanza – aprendizaje del objeto de estudio.

Teniendo esta idea, propusimos diseñar unos aplicativos y simulaciones que permitieran explorar los distintos modelos de regresión lineal y no lineal, a partir de situaciones, con el fin que los estudiantes puedan predecir frente a los datos y a los problemas abordados.

A partir de nuestra experiencia docente y de la indagación realizada en la presente propuesta, podemos afirmar que se evidencia ausencia de la estadística en los planes de estudio de matemáticas, y en consecuencia la construcción y comprensión de los modelos de regresión en la educación media; por tal motivo los aplicativos y simulaciones están diseñados de tal manera que permitan desarrollar y abordar las competencias y estándares planteados por el MEN (2006) en el Pensamiento Aleatorio y Sistemas de datos. Es así como podemos afirmar que tiene una estructura curricular acorde con las exigencias nacionales, respecto a la educación matemática y estadística en Colombia.

Por último, podemos decir que para todo proceso de indagación realizado desde su planeación, pasando por su diseño y concluyendo en su reflexión y análisis, desarrollar este tipo de propuestas permite contribuir a nuestra formación como docentes de matemáticas y enfocar nuestro perfil como docentes en educación estadística, mostrando así en este trabajo nuestro compromiso con el conocimiento matemático y estadístico, no sólo desde una perspectiva epistémica sino también desde la posible transformación de las prácticas educativas y pedagógicas en estadística, en el contexto de la educación media.

BIBLIOGRAFÍA

Arias, J. (1986). *Regresión no lineal un enfoque práctico.* Monterrey: Universidad Autónoma de Nueva León.

Batanero, C. (2001). *Didáctica de la Estadística* . Granada: Grupo de Investigación en Educación Estadística. ISBN 84-699-4295-6.

Batanero, C. &. (2000). El Papel de los Proyectos en la Enseñanza y Aprendizaje de la Estadística. *Aspectos didácticos de las matemáticas* , 125-164.

Batanero, C. (s.f.). *¿Hacia dónde va la educación estadística?* Granada: Universidad de Granada.

Batanero, Estepa & Godino. (1991). Análisis exploratorio de datos: sus posibilidades en la enseñanza secundaria. *Suma* , 25-31.

Chaves, E. (2007). Inconsistencia entre los programas de estudio y la Realidad de aula en la enseñanza de la estadística de Secundaria. *Revista Electrónica publicada por el Instituto de Investigación en Educación* , 1 - 35.

Cuevas, J. (2012). Panorama actual de los estándares educativos en estocástica. *Revista digital Matemática, Educación e Internet* , 1-12.

Fernández, F., Soler, N. y Sarmiento, B. (2008). Conocimiento estadístico y probabilístico de profesores de educación básica y media (Reporte de investigación). Bogotá: Universidad Pedagógica Nacional.

Godino, J. (1995). ¿Qué aportan los ordenadores a la enseñanza de la estadística? *UNO* , 45 - 56.

López y otros. (s.f.). *Excel como una herramienta asequible en la enseñanza de la Estadística.* Recuperado el 20 de Septiembre de 2012, de http://campus.usal.es/~teoriaeducacion/rev_numero_07/n7_art_lopez_lagunes_herrera.htm

Ministerio de Educación Nacional. (2006). *Estándares Básicos de Competencias en Matemáticas.* Bogotá.

Ministerio de Educación Nacional. (1998). *Lineamiento Curriculares de Matemáticas.* Bogotá: MEN.

Núñez, Steyerberg & Villota. (2011). Estrategias para la elaboración de modelos estadísticos de regresión. *Revista española de cardiología,, 64* (6), págs. 501-507.

Peña. (2002). *Regresión y diseño de experimentos.* España: Alianza Editorial.

Perez, M. & Carretero, M. (1990). *El papel de la instrucción en la solución de problemas de correlación.* Madrid: Universidad Autónoma de Madrid.

PISA. (2003). *Marcos teóricos de PISA 2003 : la medida de los conocimientos destrezas en matemáticas, lectura, ciencias y resolución de problemas.* Madrid: Ministerio de Educación y Ciencia, Instituto Nacional de Evaluación y Calidad del Sistema Educativo.

Rivas, G. (2004). Regresión no lineal. *Revista Colombiana de Estadística* , 89 - 102.

Rojas, M. (1990). *Técnicas de diagnóstico en Regresión Lineal.* Chile: Sociedad Chilena de Estadística.

Sanchez, F. (1999). Significado de la correlación y regresión para los estudiantes universitarios. *Tesis* . Granada, España: Universidad de Granada.

ANEXOS

ANEXO A. Planteamiento de situaciones de regresión

A continuación se proponen una serie de situaciones relacionadas con los modelos de regresión lineal y no lineal, que pueden ser trabajadas en el aula de matemáticas, con ayuda de los aplicativos y simuladores diseñados en la presente propuesta.

1. Un fabricante de pintura para demarcar canchas de futbol desarrolló una nueva fórmula y necesita probar su durabilidad. Uno de los factores es la concentración de pigmento en la pintura. Si la concentración es muy baja, la pintura se decolora rápidamente; si la concentración es muy alta, la pintura no tiene una buena adherencia a la superficie de la cancha. El fabricante aplicó pinturas de varias concentraciones en una muestra de cintas asfálticas y obtuvo una medida de la durabilidad de cada una de ellas. Los datos están almacenados en el computador, ubicados en dos columnas, en la primera la durabilidad y en la segunda la concentración.

a. Calcule, a través de los aplicativos y simuladores, la ecuación de regresión, pronosticando la durabilidad por medio de la concentración. Interpreta el coeficiente correspondiente a la pendiente
b. Encuentre el coeficiente de correlación. ¿Qué indica éste acerca del valor predictivo de la concentración?

2. Las altas tasas de interés aplicadas a los créditos hipotecarios en años recientes han tenido un impacto en el número de construcciones iniciadas en un país desarrollado. Los datos que se muestran a continuación sintetizan las tasas de interés prevalecientes por trimestre y el número correspondiente de construcciones iniciadas en este semestre en una localidad dada:

Mes	**Tasa de interés (%)** **x**	**Número de construcciones iniciadas** **y**
1	11,5	260
2	11,4	250
3	11,6	241
4	12,4	256
5	12,8	270
6	13,2	220
7	13,5	190
8	13,0	195
9	12,7	200
10	12,9	210
11	12,5	230
12	12,0	245

a. Trace la gráfica de los datos

b. Ajuste una línea de regresión lineal para estos datos. ¿Es $\hat{B}$significativamente diferente a cero?

c. Pronostique el número de construcciones iniciadas correspondiente a una tasa de interés de 11,0%. Utilice un intervalo de predicción al 95%.

3. En un estudio se reunieron datos acerca del número de habitantes y los gastos per cápita en servicios de seguridad pública de 28 poblaciones. La línea de regresión de dicho estudio tiene una pendiente positiva. No obstante, cuando una observación inusual se excluyó de los datos, una población con un centro comercial regional muy grande, la pendiente de la recta de regresión se tornó negativa. ¿Cómo afectará la exclusión de una población de la base de datos el signo del coeficiente de correlación?

4. Una compañía de consolas de video juegos de una pequeña ciudad trata de predecir la demanda de sus servicios con base en el los código que tiene cada consola. Los propietarios de la compañía obtuvieron el número de casas que tienen consolas de dicha empresa, y calcularon sus ventas por miles de casas y el número de casas por acre (un acre equivale a 4047 metros cuadrados). Los datos están organizados en una tabla de dos columnas, la primera contiene las ventas y la segunda la densidad (residencias/acre).

a. Obtenga la correlación entre las dos variables, ¿qué significa el signo?

b. Obtenga una ecuación de predicción con las ventas como variable dependiente y la densidad como variable independiente. ¿Qué interpretación le puedes dar a la ordenada en el origen y a la pendiente?

5. En una fábrica de producción de herramientas, el método para deformar acero a temperatura normal, presenta una relación con la dureza del material, puesto que a medida que la deformación aumenta, se afecta la dureza del material. Los operarios tomaron los datos de la producción de acero donde se relación la deformación con la dureza:

X:Deformación (en mm)	6	9	11	13	22	26	28	33	35
Y:Dureza (en kg/mm^2)	68	67	65	53	44	40	37	34	32

a. Identifique el modelo de regresión que mejor se ajusta a la situación.

b. Explique por qué el modelo escogido es el más adecuado para la situación

c. Si la deformación del material es de 80 mm, ¿Cuál será la dureza del material?

6. La relación de las exportaciones de un país y la producción interna del mismo es de tipo lineal. Se tienen los datos anuales de España, evaluados en miles de millones de pesetas. A continuación se presenta la información desde 1992 a 1996:

Años	**Producción**	**Exportaciones**
1992	52.654	10.420
1993	53.972	11.841
1994	57.383	14.443
1995	61.829	16.732
1996	65.381	18.760

a. Identifique el modelo de regresión que mejor se ajusta a la situación.

b. Explique por qué el modelo escogido es el más adecuado para la situación

c. Determine el dinero en exportaciones, teniendo en cuenta que la producción fue de 78.943 millones de pesetas.

7. Una empresa de artesanías está lanzando al mercado un artículo, decidió anotar periódicamente el tiempo medio (medido en minutos) que se utiliza para realizar una pieza (variable Y) y el número de días desde que empezó el proceso de fabricación (variable X). Esto se realiza con el fin de analizar la adaptación al nuevo proceso por parte de los operarios, los cuales van mejorando en el mismo a media que van adquiriendo experiencia. A continuación se presenta una tabla donde se relación el tiempo de fabricación en función al número de días:

X	10	20	30	40	50	60	70
Y	35	28	23	20	18	15	13

a. Determine el tiempo de fabricación del artículo cuando lleva 100 días

b. Determine el tiempo transcurrido hasta que el tiempo de fabricación que se prediga sea de 10 minutos

c. Determine el porcentaje de tiempo que se reduce por cada día que pasa

8. Una juguetería está analizando la evolución reciente de las ventas de su nuevo muñeco (Y, en millones de euros) junto con los gastos de publicidad de ese muñeco (X, en millones de euros). Para esto, organizaron la información en la siguiente tabla:

Año	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**
Y	126	135	156	156	150	150	180
X	20	25	30	32,5	35	32	34

a. ¿Se puede afirmar que al aumentar los gastos en publicidad se incrementarán las ventas? Obtenga los parámetros del ajuste lineal que explique las ventas en función de los gastos. Interprete dichos coeficientes e indique la bondad del ajuste realizado.

b. Si para el año 2003 aumentásemos los gastos de publicidad en un 1%, ¿en qué porcentaje se espera que variasen las ventas, según el modelo lineal?

9. Una nutricionista planteó una dieta de adelgazamiento para un grupo de pacientes, para llevar un registro estricto, ha recogido datos sobre el peso perdido desde el inicio de la dieta (variable Y, en Kg.) y el tiempo que llevan siguiendo la dieta (Variable X, en semanas), los cuales se muestran en la siguiente tabla.

Y	2,4	5,4	5,6	8,4	10,6	13,5	15	15
X	3	5	6	8	11	13	15	16

a. Estime el modelo de regresión que explica el peso perdido en función del tiempo que se lleva siguiendo la dieta. ¿Por qué ese y no otro?

b. Según el modelo considerado, ¿qué peso esperaría perder una persona que siga la dieta durante 2 meses (8 semanas)? ¿Y una persona que esté dispuesta a seguir la dieta durante dos años (108 semanas)? ¿Qué fiabilidad le otorga a cada una de las estimaciones anteriores?