



UNIVERSIDAD DE QUINTANA ROO
DIVISIÓN DE CIENCIAS E INGENIERÍA

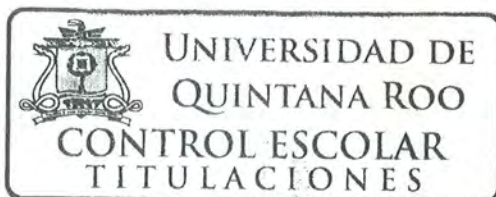
ACTIVIDADES DE APRENDIZAJE PARA REGRESIÓN LINEAL

TESIS
PARA OBTENER EL GRADO DE
MAESTRO EN ENSEÑANZA DE LAS MATEMÁTICAS

PRESENTA
JUAN DIEGO CUEVAS DOMÍNGUEZ

DIRECTOR
DR. VICTOR H. DE JESÚS SOBERANIS CRUZ

ASESORES
DR. JAIME DIONISIO CUEVAS DOMÍNGUEZ
M.E.M. WALTER MAGAÑA LANDERO
DR. JAIME SILVERIO ORTEGÓN AGUILAR
M.T.I. MELISSA BLANQUETO ESTRADA



CHETUMAL QUINTANA ROO, MÉXICO, NOVIEMBRE DE 2018



UNIVERSIDAD DE QUINTANA ROO
DIVISIÓN DE CIENCIAS E INGENIERÍA

TRABAJO DE TESIS BAJO LA SUPERVISIÓN DEL COMITÉ DEL
PROGRAMA DE MAESTRÍA Y APROBADA COMO REQUISITO
PARA OBTENER EL GRADO DE:

MAESTRO EN ENSEÑANZA DE LAS MATEMÁTICAS

COMITÉ DE TESIS

DIRECTOR:

DR. VICTOR H. DE JESÚS SOBERANIS CRUZ

ASESOR:

DR. JAIME DIONISIO CUEVAS DOMÍNGUEZ

ASESOR:

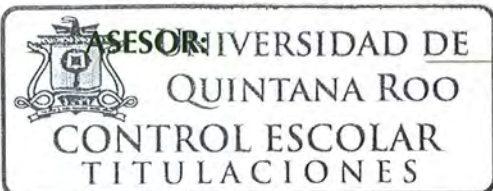
M.E.M. WALTER MAGAÑA LANDERO

ASESOR:

DR. JAIME SILVERIO ORTEGÓN AGUILAR



M.T.I. MELISSA BLANQUETO ESTRADA



INDICE GENERAL

CAPITULO 1: INTRODUCCION	7
1.2 PROBLEMA	9
1.3 OBJETIVOS	9
1.3.1 Objetivo General	9
1.3.2 Objetivos particulares	9
CAPITULO 2: MARCO TEÓRICO	11
2.1 Análisis de regresión lineal	11
2.1.1 Correlación	11
2.1.2 Gráfico de Dispersión	13
2.1.3 Modelo Lineal Simple	14
2.1.4 Supuestos del modelo lineal	14
2.2 Niveles de razonamiento (Modelo de Van Hiele)	16
2.3 Trayectoria Hipotética de Aprendizaje	18
2.4 Análisis Didáctico (enfoque de Fernández, Monroy y Rodríguez (1998))	21
2.4.1 Análisis de contenido.	21
2.4.2 Análisis de la enseñanza	21
2.4.3 Análisis de aprendizaje	22
2.5 La tecnología en el aprendizaje de las matemáticas (estadística)	22
CAPITULO 3: METODOLOGÍA	25
3.1 Actividades para la enseñanza del modelo de “REGRESIÓN LINEAL”	25
A.1 Actividad 1.	26

A.2 Actividad 2.	27
A.3 Actividad 3.	30
A.4 Actividad 4.	31
A.5 Actividad 5.	32
A.6 Actividad 6.	33
A.7 Actividad 7.	35
CAPITULO 4: RESULTADOS	36
Sesión 1	37
Sesión 2	40
Sesión 3	46
CAPITULO 5: CONCLUSIONES Y RECOMENDACIONES	50
Conclusiones	50
5.1 La propuesta didáctica	50
5.2 Aprendizaje de los estudiantes	50
5.3 Recomendaciones	51
REFERENCIAS BIBLIOGRÁFICAS	52
ANEXO A	54

INDICE DE TABLAS

Tabla 1. Observaciones a nuestra disposición.	11
Tabla 2. Datos para 20 países en los que se midieron dos variables	12
Tabla 3. Datos sobre el promedio de uso semanal y gastos anuales de mantenimiento.	27
Tabla 4. American's Best College, edición del año 2000	29
Tabla 5. American's Best College, edición del año 2000	55

INDICE DE FIGURAS

Figura 1. Tasa de mortalidad en niños menores a 5 años.	13
Figura 2. Análisis didáctico.	20
Figura 3. Gráficos de dispersión para la actividad tres.	31
Figura 4. Gráficos de dispersión para la actividad cuatro.	32
Figura 5. Gráfica de puntos elaborada por estudiante 1.	39
Figura 6. Gráfico de dispersión elaborado en grupo en hoja de cálculo de Excel	40
Figura 6a. Gráfico elaborado individualmente por estudiante 1.	41
Figura 6b. Gráfico elaborado individualmente por estudiante 1.	42
Figura 6c. Gráfico elaborado individualmente por estudiante 1.	43
Figura 6d. Gráfico elaborado individualmente por estudiante 2.	43
Figura 6e. Gráfico elaborado individualmente por estudiante 2.	44
Figura 7. Actividad 3 elaborada por estudiante 1.	46
Figura 8. Actividad 4 elaborada por estudiante 1.	47
Figura 9. Actividad 4 elaborada por estudiante 2.	48
Figura 10. Actividad 5 elaborada por estudiante 1.	49
Figura 11. Actividad 5 elaborada por estudiante 2.	49

RESUMEN

Esta tesis denominada “Actividades de aprendizaje para regresión lineal”, documenta el criterio de diseño de una propuesta didáctica para la enseñanza-aprendizaje de los conocimientos básicos que deben tener los estudiantes para un estudio amplio y sólido del modelo de “regresión lineal”, también documenta la experiencia al implementar la secuencia didáctica y los resultados obtenidos.

Las actividades seleccionadas para diseñar la propuesta didáctica se sustentan en el modelo de análisis didáctico propuesto por Pedro Gómez (2002), el interés se da en que los estudiantes desarrollen el conocimiento de manera intuitiva con el manejo de diferentes registros de representación (tratamiento, conversión) de datos y su interpretación.

El primer capítulo describe, a manera de introducción al tema, la importancia de la enseñanza de la estadística en el programa universitario, en especial el modelo de regresión lineal, sus aplicaciones en el ámbito profesional y la vida cotidiana, la integración de la computadora como herramienta en el proceso de enseñanza-aprendizaje. También describe el problema planteado a resolver y su justificación, los objetivos de la tesis, el alcance y sus limitaciones.

El capítulo dos se refiere al marco teórico en el que se basó y sustenta este trabajo de tesis, se comienza con una breve relatoría en la historia de los primeros trabajos, en el siglo XIX, que conciernen con el estudio de la regresión lineal con el aporte de Sir Francis Galton (1822-1917). Continúa con una explicación de los conceptos matemáticos que estarán en juego, para después seguir con la descripción de los modelos de enseñanza- aprendizaje que se mencionan en orden histórico y que tuvieron influencia para el diseño de la secuencia didáctica de esta tesis.

El tercer capítulo describe el objetivo y las características de las actividades de la propuesta didáctica, para cada actividad se presenta una discusión del análisis de contenido, análisis de enseñanza y análisis de aprendizaje.

El cuarto capítulo se centra en documentar el análisis y los resultados obtenidos en la experiencia de implementar las actividades de la secuencia didáctica propuesta, poniendo especial atención en el proceso, en los estudiantes, del desarrollo del conocimiento de los conceptos básicos de la regresión lineal, cual fue el papel del profesor como guía facilitador para apoyar el desarrollo del conocimiento, como fue el ambiente en el aula y como fue la respuesta a las dificultades en los estudiantes.

En el último capítulo se enuncian las conclusiones de la tesis desde el enfoque del desarrollo de la propuesta didáctica, los objetivos planteados y el alcance en el aprendizaje de los estudiantes, también se sugieren algunas recomendaciones para el caso en que algún docente desee implementar con sus alumnos la propuesta didáctica.

CAPITULO 1

INTRODUCCION

Aunque hace algunos años pocos investigadores se interesaban por los problemas de enseñanza y aprendizaje de la estadística, en la actualidad existe un aumento notable de las publicaciones, diseños curriculares e investigación relacionados con este tema. La estadística se ha incorporado, en forma generalizada al currículum de la enseñanza de las matemáticas en las diferentes especialidades universitarias de nuestro país.

1.1 ANTECEDENTES

El análisis de regresión lineal es una técnica estadística utilizada para estudiar la relación entre variables. Se adapta a una amplia variedad de situaciones reales, por ejemplo, en la investigación social, el análisis de regresión lineal se utiliza para predecir un amplio rango de fenómenos, desde medidas económicas hasta diferentes aspectos del comportamiento humano. En el contexto de la investigación de mercados puede utilizarse para determinar en cual de diferentes medios de comunicación puede resultar más eficaz invertir o para predecir el número de ventas de un determinado producto. En física se utiliza para caracterizar la relación entre variables o para calibrar medidas etc.

Tanto en el caso de dos variables (Regresión Lineal Simple) como en el caso de más de dos variables (Regresión Lineal Múltiple), el análisis de regresión lineal puede utilizarse para explorar y cuantificar la relación entre una variable llamada dependiente o de “respuesta” y una o más variables llamadas independientes o “predictoras”, así como para desarrollar una ecuación lineal con fines predictivos. Además, el análisis de regresión lleva asociados una serie de procedimientos de diagnóstico como el análisis de residuos y puntos de influencia que informan sobre la estabilidad e idoneidad del análisis y que proporcionan pistas sobre como perfeccionarlo.

En la concepción clásica de la estadística, la noción de aleatoriedad ha estado ligada a las diferentes concepciones sobre la probabilidad (Godino, Batanero y Cañizares, 1987). En una *concepción clásica*, la probabilidad de un suceso es el *cociente entre el número de casos favorables al suceso y el número de casos posibles siempre que todos sean igual-probables*.

Por otra parte, una heurística es una estrategia inconsciente que reduce la complejidad de un problema probabilístico, suprimiendo parte de la información. Aunque las heurísticas ayudan en muchos casos a obtener una solución aproximada al problema, en otros produce sesgos en las conclusiones obtenidas, con las consiguientes implicaciones en las decisiones tomadas. Por ejemplo, el siguiente ítem adaptado de Kahneman y cols. (1982).

Ejemplo. En un hospital maternal se lleva un registro del sexo de los recién nacidos. ¿Cuál de los dos sucesos siguientes te parece que tiene más probabilidad?

- A. Que entre los próximos 10 recién nacidos haya más de un 70% de niñas.*
- B. Que entre los próximos 100 niños recién nacidos haya más de un 70% de niñas.*
- C. Las dos cosas me parecen igual de probables*

La respuesta correcta a este ítem es la A, aunque la mayoría de las personas suelen considerar correcta la respuesta C. Esto es a que solo tienen en cuenta que la proporción de niñas en las dos muestras es la misma, sin tener en cuenta el tamaño de las muestras. Al suprimir un dato (el tamaño de la muestra) se ha reducido la complejidad del problema, pero se ha producido una solución incorrecta.

Son muchas las definiciones de estadística, entre ellas la siguiente refleja bien la concepción del tema:

“La estadística estudia el comportamiento del fenómeno llamados de colectivo. Está caracterizada por una información acerca de un colectivo o universo, lo que constituye su objeto material y unas previsiones de cara al futuro, lo que implica un ambiente de incertidumbre, que constituyen su objeto o causa final” (Cabriá, 1994).

Se está caminando hacia una sociedad cada vez más informatizada y una comprensión de las técnicas básicas de análisis de datos y su interpretación adecuada son cada día más importantes (Batanero, 2001). Al mismo tiempo, la estadística como ciencia, atraviesa un período de notable expansión, siendo cada vez más numerosos los procedimientos disponibles, alejándose cada vez más de la matemática pura, lo que implica la dificultad de enseñar un tema en continuo cambio y crecimiento.

Por otra parte, los avances en la computación han eliminado la mayor parte del trabajo pesado computacional del análisis en el tema de la estadística y en particular en el tema de regresión lineal, lo que permite, por ejemplo, un análisis gráfico más extenso. Si bien estos avances han facilitado el análisis de los datos, se ha prestado menos atención a la computadora como una herramienta para la enseñanza. En este trabajo de tesis la secuencia didáctica propuesta incluye el uso de la computadora para que el equipo, tanto computacionalmente como gráficamente pueda ayudar a desarrollar en los estudiantes una idea de los conceptos básicos de regresión lineal como error aleatorio, varianza, distinguir valores atípicos, observaciones influyentes y la interpretación de los coeficientes de regresión lineal simple. Un gran énfasis estará en la computadora como una herramienta para la enseñanza.

1.2 PROBLEMA

Al reflexionar sobre como diseñar una secuencia didáctica para desarrollar el concepto de Modelo de Regresión Lineal y sus conceptos asociados en alumnos de nuevo ingreso en la universidad quienes no tienen una formación suficientemente sólida en los métodos y técnicas de esta materia, surge la necesidad de concretar el conocimiento que esperamos desarrollen los estudiantes. ¿Cuál es este conocimiento? ¿Cómo resumirlo y hacerlo útil e interesante para nuestros estudiantes? ¿Qué tipo de situaciones didácticas podemos usar para la enseñanza, si se quiere hacerlo consecuente con los principios constructivistas del aprendizaje, con la importancia de la interacción social y del trabajo en grupo del estudiante?

En la mayoría de los libros de texto de Regresión Lineal disponibles manejan un nivel matemático avanzado y existen pocos precedentes de texto dedicado a la formación didáctica de los profesores en esta área particular. Esto motivó a diseñar una secuencia didáctica apoyada en el modelo de análisis didáctico propuesto por Pedro Gómez (2002) y en la propuesta de fundamentación para la enseñanza de la geometría llamada “El Modelo de Van Hiele” (Jaime A., y Gutiérrez A. 1990).

1.3 OBJETIVOS

El presente trabajo de tesis tiene como objetivo principal documentar la experiencia en el diseño de una secuencia didáctica de aprendizaje para que el estudiante desarrolle conocimientos para la interpretación del modelo de regresión lineal simple, así como documentar las actividades y desempeño de los estudiantes durante la implementación de la secuencia didáctica y evaluar los resultados obtenidos al final en referencia a los objetivos planteados para la secuencia didáctica propuesta.

1.3.1 Objetivo General

Proponer una secuencia de aprendizaje para que el estudiante desarrolle conocimientos y habilidades de los aspectos principales que intervienen en un entendimiento básico del análisis de regresión lineal y sus conceptos asociados incluyendo, la interpretación del modelo de regresión lineal simple en términos de aspectos que caracterizan al “fenómeno” de la vida real a partir del cual se obtuvieron los datos, en un ambiente donde los estudiantes interpreten diversas representaciones (gráfica, tabular, algebraica), que argumenten y describan situaciones en un ambiente colaborativo.

1.3.2 Objetivos particulares

Determinar las características de cada uno de los niveles de comprensión del análisis de regresión lineal simple.

Desarrollar la evaluación de cada uno de los niveles de comprensión del análisis de regresión lineal simple.

Interpretar el modelo de regresión lineal simple en términos de aspectos que caracterizan al fenómeno de la vida real.

CAPITULO 2

MARCO TEÓRICO

La regresión y la correlación lineal son contenidos de estadística que ya forman parte del currículum de enseñanza universitaria. Su inclusión posibilita el tratamiento estadístico de datos como, con dos variables y el análisis de la relación que puede existir entre las variables consideradas. Los primeros trabajos que conciernen con el estudio de la regresión se remontan al siglo XIX, cuando Sir Francis Galton (1822-1917) imbricó sus dos grandes aficiones: el estudio de la herencia y la expresión matemática de los fenómenos vinculados a ella. Él fue el primero en trabajar con un conjunto de variables y asignar a la relación entre dos variables un número para así obtener una medida tocante a su grado de relación. Sostenía la idea de que personas excepcionalmente altas solían tener hijos de estatura menor, mientras que personas muy bajas solían tener hijos más altos; este hecho fue enunciado por Galton como la regresión a la media, aplicable a las tallas de una generación respecto de las siguientes. La justificación que se le da hoy día a esta situación es que los valores extremos de una distribución se deben en gran medida al azar.

2.1 Análisis de regresión lineal

Se mencionó que el análisis de regresión lineal es una técnica estadística utilizada para estudiar la relación entre variables. El interés en este trabajo de tesis se centra en proporcionar los fundamentos del análisis de regresión sin hacer hincapié en los aspectos más técnicos del análisis de regresión, sino que se intenta fomentar la comprensión de cuándo y cómo utilizar el análisis de regresión lineal y cómo interpretar sus resultados.

2.1.1 Correlación

Dado que la regresión lineal se ocupa de investigar la relación entre dos o más variables continuas, al parecer es correcto tratar de describir el vínculo observado en una serie de observaciones resumidas en una tabla y que la conclusión se puede resumir en un valor numérico, como se describe en la tabla 1

Tabla 1 Observaciones a nuestra disposición, aquí X_1 quiere decir la variable X medida en el individuo 1, etc.

<u>Individuo</u>	<u>Variable X</u>	<u>Variable Y</u>
1	X_1	Y_1
2	X_2	Y_2
.	.	.
n	X_n	Y_n

En este ejemplo se considera que se miden ambas variables en la misma unidad: puede tratarse de un individuo, un paciente, un país, un animal, una escuela, etc.

Observar un ejemplo tomado de, apunte de regresión lineal (M. Szretter, 2013)

Ejemplo. - Queremos investigar la relación entre el porcentaje de niños que ha sido vacunado contra tres enfermedades infecciosas: difteria, tosferina y tétanos, (DPT que se suele denominar triple bacteriana) en un cierto país y la correspondiente tasa de mortalidad infantil para niños menores a cinco años. El Fondo de las Naciones Unidas para la Infancia considera a la tasa de mortalidad infantil para niños menores a cinco años como uno de los indicadores más importantes del nivel de bienestar de su población infantil. Datos publicados en United Nations Children’s Fund. **The State of the World’s Children 1994.**

Tabla 2: Datos para 20 países en los que se midieron dos variables, X: porcentaje de niños vacunados a la edad de un año en cada país, Y: es la tasa de mortalidad infantil de niños menores a cinco años por cada 1000 niños nacidos vivos en cada país.

País	Porcentaje vacunado	Tasa de mortalidad menor a 5 años
Bolivia	77	118
Brasil	69	65
Camboya	32	184
Canadá	85	8
China	94	43
República Checa	99	12
Egipto	89	55
Etiopía	13	208
Finlandia	95	7
Francia	95	9
Grecia	54	9
India	89	124
Italia	95	10
Japón	87	6
México	91	33
Polonia	98	16
Federación Rusa	73	32
Senegal	47	145
Turquía	76	87
Reino Unido	90	9

En la tabla 2 se muestran los datos para 20 países en el año 1992, X representa el porcentaje de niños vacunados, Y representa la tasa de mortalidad infantil de niños menores de 5 años por cada 1000 niños nacidos vivos, tenemos una pareja de resultados (X_i, Y_i) para cada país en la muestra. Para leer la información desplegada en la tabla es, por ejemplo, para Bolivia $X_1 = 77.0$, es decir, en el año 1992, un 77% de los niños menores de un año estaban vacunados contra la DPT y (en el año 1992) 118 menores de 5 años murieron por cada 1000 niños nacidos vivos. La correcta interpretación de la tabla se considera muy importante para una buena comprensión del tema.

2.1.2 Gráfico de Dispersión

Otra forma de visualizar esta información es mediante un gráfico de dispersión. En un gráfico de dispersión se ubican los resultados de una variable (X) en el eje horizontal y los de la otra variable (Y) en el eje vertical. Cada punto en el gráfico representa una observación (X_i, Y_i) .

En este tipo de gráfico se pierde la información del individuo (paciente o país), y aunque si hubiera pocos puntos se los podría rotular, esencialmente esta información no suele estar disponible en un gráfico de dispersión. Es deseable que el alumno sepa interpretar y trasladar la información de una representación a otra (Duval, 2006).

En la figura 1 puede verse los datos de la tabla 2 representados en un gráfico de dispersión. Ahí vemos que, por ejemplo, Bolivia está representada por el punto (77, 118).



Figural.- Tasa de mortalidad en niños menores a 5 años

Usualmente con este gráfico se puede determinar si existe algún tipo de relación entre X e Y . Para este caso se observa que a medida que aumenta el porcentaje de niños inmunizados, decrece la tasa de

mortalidad. Para el objetivo de este trabajo se está interesado en situaciones donde aparece este tipo de relación. Sin embargo, a manera de ilustración se puede mencionar otras posibilidades que se pueden observar al representar los datos en un gráfico de dispersión como puede ser: ausencia de datos en un intervalo, no asociación de datos, vínculo curvilíneo, agrupamientos de datos. La correcta interpretación del gráfico y su relación con la tabla correspondiente es considerada muy importante para una buena comprensión del tema.

2.1.3 Modelo Lineal Simple

El modelo de regresión lineal simple es un modelo para el vínculo de dos variables aleatorias que comúnmente se denominan $X = \text{variable predictora}$ e $Y = \text{variable dependiente o de respuesta}$. El modelo lineal (simple pues solo vincula una variable predictora con una variable de respuesta) propone que:

$$Y = \beta_0 + \beta_1 X + \epsilon, \quad (1)$$

Donde ϵ es el término del error. Esto es que para cada valor de X , la correspondiente observación Y consiste en el valor $\beta_0 + \beta_1 X$, más una cantidad ϵ , que puede ser positiva o negativa, y que da cuenta de que la relación entre X e Y no es exactamente lineal, sino que está expuesta a variaciones individuales que hacen que el par observado (X, Y) no caiga exactamente sobre la recta, sino cerca de ella, como puede anticiparse viendo el gráfico de dispersión de los datos que usualmente se modelan con este modelo (ver ejemplo de fig. 1). En el modelo (1) los números β_0 y β_1 son constantes desconocidas que se denominan *parámetros* del modelo o *coeficientes* de la ecuación. El modelo se denomina “lineal” puesto que la Y depende linealmente de estas constantes, es lineal en los parámetros: los β 's no aparecen como exponentes ni multiplicados o divididos por otros parámetros. Los parámetros se denominan:

$\beta_0 =$ ordenada al origen

$\beta_1 =$ pendiente

Otra forma de escribir el mismo modelo es pensando en las observaciones (X_i, Y_i) . En tal caso el modelo (1) adopta la forma:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \quad (2)$$

Donde ϵ_i es el término del error (que no es observable) para el individuo i ésimo.

2.1.4 Supuestos del modelo lineal

Tomando en cuenta el ejemplo mostrado en el gráfico de la fig. 1, se puede observar que la mayoría de los puntos no caen exactamente sobre una línea. La dispersión de los puntos alrededor de cualquier línea que se dibuje representa la variación en la tasa de mortalidad en niños menores a 5 años que no está asociado con el porcentaje de niños vacunados, y que usualmente se considera que es de

naturaleza aleatoria. Muchas veces esta aleatoriedad se debe a la falta de información adicional y de un modelo complejo que pueda dar un adecuado vínculo funcional entre los diferentes datos que puedan influir y la variable respuesta. Por otro lado como se espera que todos esos componentes diversos (que puedan influir) se sumen entre sí y tengan un aporte muy menor a la explicación de la variable respuesta comparada con el de la explicativa considerada, se los puede modelar adecuadamente asumiendo que todas estas características independientes del porcentaje de niños vacunados y asociadas al individuo (en este caso país) están incluidas en el término del error, que al ser suma de muchas pequeñas variables independientes (y no relevadas) se puede asumir que tienen distribución normal. Lo cual no se aleja mucho de la realidad en muchos de los ejemplos prácticos de aplicación del modelo de regresión.

Es por eso que el modelo lineal

$$Y = \beta_0 + \beta_1 X + \epsilon,$$

Está formado de una parte aleatoria ϵ y otra determinística, $\beta_0 + \beta_1 X$.

Los supuestos bajo los cuales serán válidas las inferencias que haremos sobre el modelo

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i,$$

Son los siguientes:

- 1.- Los ϵ_i tienen media cero, llamada $E(\epsilon_i) = 0$
- 2.- Los ϵ_i tienen toda la misma varianza desconocida denominada σ^2 y que el otro parámetro del modelo es la varianza, denominado $Var(\epsilon_i) = \sigma^2$
- 3.- Los ϵ_i tienen distribución normal.
- 4.- Los ϵ_i son independientes entre sí y son no correlacionados con las X_i .

El hecho de que los errores no estén correlacionados con las variables explicativas apunta a que el modelo esté identificado. Otra manera de escribir los supuestos es observar que a partir de la ecuación (2) o (1) se puede observar que **para cada valor fijo de la variable X** , el valor esperado de la respuesta Y depende de X de manera lineal, es decir podemos escribir el modelo en términos de la esperanza de Y condicional a las X 's cuya notación es $E(Y | X)$. Esto constituye un modo equivalente de escribir el modelo de regresión lineal simple. Supuestos:

- 1.- La esperanza condicional de Y depende de X de manera lineal, es decir

$$E(Y | X) = \beta_0 + \beta_1 X \quad (3)$$

O, escrito de otro modo

$$E(Y | X = x_i) = \beta_0 + \beta_1 x_i$$

Donde β_0, β_1 son los parámetros del modelo, o coeficientes de la ecuación. A la ecuación (3) se le suele llamar función de respuesta, es una recta.

2.- La varianza de la variable respuesta Y dado que la predictora está fijada en $X = x$ se denota por $Var(Y | X = x)$. Se asume que satisface

$$Var(Y | X = x_i) = \sigma^2$$

O sea, es constante (una constante desconocida y positiva) y no depende del valor de X .

3.- Las Y_i , es decir el valor de la variable Y cuando X toma el valor i ésimo observado.

$Y_i = Y | X = x_i$ tienen distribución normal.

4.- Las Y_i son independientes entre sí.

Si para algún conjunto de datos estos supuestos no se cumplen (por ejemplo, las observaciones no son independientes porque hay varias mediciones de los mismos pacientes, o la varianza de Y crece a medida que crece X) no se puede aplicar el modelo de regresión lineal a dichos datos. Es necesario trabajar con modelos más refinados, que permitan incluir estas estructuras en los datos.

El modelo de regresión lineal tiene tres parámetros a ser estimados, β_0, β_1 y σ^2 . Para esta propuesta didáctica ¿que interesa resolver?

1.- Estimar los parámetros a partir de observaciones.

2.- Dar alguna medida de la adecuación del modelo a los datos.

3.- Estimar la esperanza condicional de Y para algún valor de X observado o para algún valor de X que no haya sido observado en la muestra, no llegaremos al grado de construir un intervalo de confianza para dicha esperanza.

4.- Describir los alcances y los problemas del modelo de regresión lineal.

2.2 Niveles de razonamiento (Modelo de Van Hiele)

El hecho de que la base de los modelos educativos se encuentre en el estudio del comportamiento del individuo introduce una gran cantidad de factores aleatorios que es muy difícil o imposible, poder controlar. Por lo tanto, como profesor no hay que pensar que aplicando fielmente un determinado modelo educativo (el de Van Hiele o cualquier otro) se resolverán todos los problemas y que todos los alumnos comprenderán y aprenderán matemáticas sin esfuerzo. Más bien hay que esperar que, aunque el modelo no produzca resultados perfectos, si proporcione mejores resultados que otras formas de trabajo. En este sentido el modelo de Van Hiele es un excelente guía para este trabajo de tesis propuesto pues hay que considerar que nos enseña a descubrir cómo debe comunicarse el profesor con los alumnos, para

presentarles los nuevos conceptos de manera que se fomente la comprensión de las matemáticas su aprendizaje y el desarrollo de la capacidad de razonamiento de los estudiantes. Van Hiele (1957) explica brevemente cual fue su primer intento de solución, mediante la elaboración de un modelo educativo que trata de explicar el porqué del comportamiento de sus alumnos

“Puede decirse que alguien ha alcanzado un nivel superior de pensamiento cuando un nuevo orden de pensamiento le permite, con respecto a ciertas operaciones, aplicar estas operaciones a nuevos objetos. El alcance del nuevo nivel no se puede conseguir por enseñanza, pero, aun así, mediante una adecuada elección de ejercicios, el profesor puede crear una situación favorable para que el alumno alcance nivel superior de pensamiento” (Jaime A. y Gutiérrez A., 1990).

En este trabajo de tesis se consideran algunas ideas del modelo de razonamiento y aprendizaje de las matemáticas elaborado por dos investigadores y profesores holandeses llamados Pierre Marie Van Hiele y Dina Van Hiele-Geldof. Las ideas centrales que siguen siendo el corazón del Modelo (educativo) de Van Hiele tal como se utiliza actualmente, pueden enunciarse de la siguiente manera:

- Se pueden encontrar varios niveles diferentes de perfección en el razonamiento de los estudiantes de matemáticas.
- Un estudiante solo podrá comprender realmente aquella parte de las matemáticas que el profesor le presente de manera adecuada a su nivel de razonamiento.
- Si una relación matemática no puede ser expresada en el nivel actual de razonamiento de los estudiantes, será necesario esperar a que éstos alcancen un nivel de razonamiento superior para presentársela.
- No se puede enseñar a una persona a razonar de una determinada forma. Pero si se le puede ayudar, mediante una enseñanza adecuada de las matemáticas, a que llegue lo antes posible a razonar de esa forma.

El modelo de Van Hiele está formado, realmente, por dos partes. La primera de ellas es descriptiva, ya que identifica una secuencia de tipos de razonamiento, llamados los “niveles de razonamiento”, a través de los cuales progresa la capacidad de razonamiento matemático de los estudiantes. La otra parte del modelo da a los profesores directrices sobre cómo pueden ayudar a sus alumnos para que puedan alcanzar con más facilidad un nivel superior de razonamiento, estas directrices se conocen con el nombre de “fases de aprendizaje” (Jaime A. y Gutiérrez A., 1990).

Los cuatro niveles de razonamiento matemático del modelo de Van Hiele son, Nivel 1 de reconocimiento. Nivel 2 de análisis, Nivel 3 de clasificación, Nivel 4 de deducción formal.

Las fases de aprendizaje son unas etapas en la graduación y organización de las actividades que debe realizar un estudiante para adquirir las experiencias que le lleven al nivel superior de razonamiento. Fase 1 de información, fase 2 orientación dirigida, fase 3 explicitación, fase 4 orientación libre, fase 5 integración.

“Si un profesor quiere basarse en el modelo de Van Hiele para la organización y el desarrollo de sus clases, deberá en primer lugar, tener una idea del nivel de razonamiento de sus alumnos” (Jaime A. y

Gutiérrez A., 1990). Uno de los métodos para determinar el nivel de razonamiento de los alumnos consiste en la realización de entrevistas individuales entre el profesor y cada estudiante, durante las cuales el profesor plantea diversas actividades y dialoga con el alumno a tenor de su forma de resolverlas y del nivel de razonamiento que vaya mostrando durante la entrevista, lo más importante no es evaluar si los estudiantes contestan bien o mal, sino como contestan y porque lo hacen así.

“Siempre que una persona empieza a estudiar un área de las matemáticas que le era desconocida hasta entonces, empieza a recorrer los niveles de razonamiento de Van Hiele empezando por el nivel 1...” (Jaime A. y Gutiérrez A., 1990).

Por último, reflexionando sobre el proceso completo de desarrollo de la capacidad de razonamiento, una cuestión importante es ¿Hasta dónde podrán progresar mis alumnos durante el tiempo que le voy a dedicar a la implementación de la propuesta de las actividades didácticas? Hay algunas consideraciones a tomar en cuenta:

- “En general el proceso de desarrollo de razonamiento no puede enmarcarse en los límites de un curso escolar. La adquisición de los niveles superiores, en particular del 3 y 4, suele ser un proceso de varios años, por lo que no es de extrañar que al terminar el curso los estudiantes sigan estando en el mismo nivel que al principio, si bien estarán más cerca de poder lograr el nivel superior” (Jaime A. y Gutiérrez A., 1990).

- “También puede ocurrir que a lo largo del curso los estudiantes alcancen un nivel, por lo que el profesor deberá empezar el trabajo que conduce al nivel siguiente. En este sentido, hay que tener en cuenta que los niveles no plantean rupturas en el proceso de aprendizaje, por lo que una vez completado el trabajo de la última fase de un nivel, se debe iniciar el trabajo de la primera fase del nivel siguiente.” (Jaime A. y Gutiérrez A., 1990).

2.3 Trayectoria Hipotética de Aprendizaje

El “ciclo de enseñanza de las matemáticas” propuesto por Simón (1995), problema del diseño curricular a nivel local en matemáticas, el cual centra su atención en la problemática de la planificación local y reflexiona sobre cómo debería ser la enseñanza, si se asume una posición constructivista social del aprendizaje de los escolares. Este modelo, llamado por Simón “el *ciclo de enseñanza de las matemáticas*”, es un “modelo esquemático de la interrelación de aspectos del conocimiento, pensamiento, toma de decisiones y actuaciones del profesor”

En resumen, los principales criterios tomados en cuenta del modelo para establecer la secuencia de trabajo son las siguientes:

- el pensamiento de los estudiantes juega un papel central
- el conocimiento del profesor evoluciona permanentemente
- la planificación para la enseñanza incluye la generación de una **trayectoria hipotética de aprendizaje**
- el cambio continuo en el conocimiento del profesor produce un cambio continuo en la trayectoria hipotética de aprendizaje
- el modelo es local en el sentido de que se centra en la enseñanza de un tópico específico para una sesión de clase.

La noción de trayectoria hipotética de aprendizaje no deja claro el papel que esta noción puede jugar en la formación inicial de profesores de matemáticas. Pedro Gómez (2007) propone una adaptación de la misma con la que el futuro profesor de matemáticas puede recoger y organizar información para el diseño de unidades didácticas.

Una trayectoria hipotética de aprendizaje (THA) consiste en establecer los objetivos para el aprendizaje de los estudiantes, las tareas matemáticas que se usarán para promover el aprendizaje de los estudiantes, y las hipótesis acerca del proceso de aprendizaje de los estudiantes (Simón, 1995)

Este constructo se fundamenta en los siguientes supuestos:

1. La construcción de una trayectoria hipotética de aprendizaje se basa en la comprensión del conocimiento actual de los estudiantes que recibirán la instrucción.
2. Una trayectoria hipotética de aprendizaje es el vehículo para planificar el aprendizaje de unos conceptos matemáticos concretos.
3. Las tareas matemáticas proporcionan las herramientas para promover el aprendizaje de unos conceptos matemáticos concretos y, por lo tanto, son un elemento clave del proceso de instrucción.
4. Dada la naturaleza hipotética e inherentemente incierta de este proceso, el profesor se verá obligado a modificar sistemáticamente cada aspecto de la trayectoria hipotética de aprendizaje.

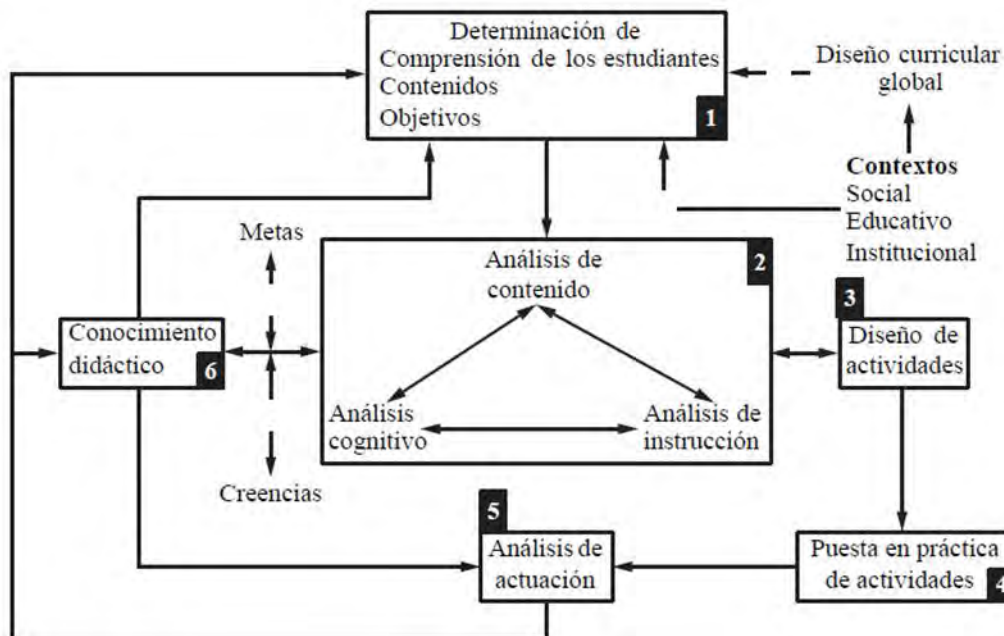


Figura 2.- Ciclo de análisis didáctico, (Revista EMA (2002), Vol. 7, N° 3, pag. 258)

El análisis didáctico se inicia con la determinación del contenido que se va a tratar y de los objetivos que se quieren lograr, teniendo en cuenta los contextos sociales, educativos e institucionales en los que se enmarca la instrucción. La información que surge del análisis de contenido sustenta el análisis cognitivo. A su vez, la realización del análisis cognitivo puede dar lugar a la revisión del análisis de contenido. Esta relación también se establece con el análisis de instrucción.

La selección de tareas que componen las actividades debe ser coherente con los resultados de los tres análisis.

Si, como argumenta Steffe (2004), “deben ser los profesores, al participar directamente en las actividades de construcción de los estudiantes, quienes deben producir sus trayectorias hipotéticas de aprendizaje” (p. 115), Sin embargo, para los futuros profesores con poca experiencia docente y, en general, sin acceso a la práctica en aulas de matemáticas, no pueden “participar directamente en las actividades de construcción de los estudiantes”

Para el diseño de estas unidades didácticas, los futuros profesores pueden poner en juego un procedimiento, que se denomina *análisis didáctico* (Gómez, 2007). El análisis didáctico es un procedimiento cíclico que describe cómo el profesor debería idealmente diseñar, llevar a la práctica y evaluar actividades de enseñanza y aprendizaje, teniendo en cuenta las condiciones en las que se realiza la formación inicial esta descripción del progreso de los escolares debe fundamentarse en la identificación, descripción y relación de cinco elementos:

1. las capacidades que los escolares tienen antes de la instrucción⁴;

2. las capacidades que se espera que los escolares desarrollen con motivo de la instrucción;
3. las tareas que conforman la instrucción;
4. las dificultades que los escolares pueden encontrar al abordar estas tareas; y
5. las hipótesis sobre los caminos por los que se puede desarrollar el aprendizaje.

La noción de capacidad es “el conjunto de condiciones necesarias para llevar a cabo una actividad concreta” Dorsch (1985)

Afirmaremos que un individuo ha desarrollado una cierta capacidad cuando él puede resolver tareas que la requieran. Por lo tanto, las capacidades:

- son específicas a un tema concreto;
- pueden incluir o involucrar otras capacidades; y
- están vinculadas a tipos de tareas

2.4 Análisis Didáctico (enfoque de Fernández, Monroy y Rodríguez (1998))

2.4.1 Análisis de contenido.

Para realizar la descripción del contenido, se consideran dos perspectivas, la conceptual y la procedimental, distinción que ofrece una manera de interpretar los procesos de aprendizaje. Rico (1997) distingue tres niveles de conocimientos conceptuales: los hechos, los conceptos propiamente dichos y las estructuras conceptuales.

El conocimiento procedimental consiste en los modos de ejecución ordenados de una tarea, los procedimientos son aquellas formas de actuación o ejecución de tareas matemáticas; en ellas es posible distinguir tres niveles: las destrezas, los razonamientos y las estrategias (Rico, 1997).

La utilización de sistemas de representación es una actividad habitual de la matemática y en particular de la estadística, sus principales manifestaciones son las descripciones verbales, tablas, gráficas y fórmulas.

2.4.2 Análisis de la enseñanza

En el análisis de la enseñanza se propone una revisión de los textos utilizados como material didáctico de apoyo. Aquí, el libro de texto es una autoridad del conocimiento y guía del aprendizaje; de igual manera, representa una ayuda inestimable para el profesor en el trabajo áulico. Asimismo, proporciona un conocimiento del significado fijado en una institución didáctica para cierto concepto. Para ello, es conveniente estudiar por separado sus componentes: enfoque de presentación del tema;

nivel de profundidad; temas abarcados; deducción de fórmulas; actividades propuestas: ejercicios y problemas; uso de las computadoras.

2.4.3 Análisis de aprendizaje

Con respecto al análisis de aprendizaje, se propone llevar a cabo una descripción de los conocimientos previos que se sugiere que los alumnos tengan al comenzar el tratamiento de estos contenidos, así como de los posibles errores y dificultades que aparecen en la resolución de problemas de regresión y correlación lineal. A partir de su conocimiento es posible diseñar actividades didácticas útiles para el profesor, que sean adecuadas para superar las dificultades y favorezcan la comprensión.

2.5 La tecnología en el aprendizaje de las matemáticas (estadística)

El uso de la tecnología ha estado influyendo tanto en la forma de hacer o desarrollar matemáticas como en la forma de aprender esta disciplina, en la actualidad existen programas o software que pueden ayudar al estudiante a explorar y desarrollar el potencial de conjeturas y resultados matemáticos. Hace algunos años cuando la computadora empezó a vislumbrarse como un instrumento importante capaz de realizar diversas operaciones de manera eficiente y rápida, también se empezó a vislumbrar su potencial en la educación, Santos (2000) señala;

“Es importante aceptar que los avances de la tecnología han sido tan rápidos y sustanciales en los últimos años que es imperativo que el estudiante adquiera habilidades y estrategias que le permitan constantemente ajustarse a estos avances y cambios. Es decir, la formación o educación del estudiante debe contemplar aspectos que le ayuden a conocer el potencial y uso de los avances tecnológicos e incorporarlos naturalmente a su práctica”

La tecnología es fundamental en la enseñanza y aprendizaje de las matemáticas, influye en las matemáticas que se enseñan y enriquece su aprendizaje, en este sentido la NCTM (2000) señala;

“...las computadoras proporcionan imágenes visuales de ideas matemáticas, facilitan la organización y el análisis de datos y hacen cálculos con eficiencia y exactitud” y agrega

“La tecnología no debería utilizarse como sustituto de los conocimientos e intuiciones básicas, sino que puede y debería usarse para potenciarlos, en la enseñanza de las matemáticas, la tecnología debería utilizarse amplia y responsablemente, con el objeto de enriquecer el aprendizaje”

En las aulas de matemáticas que se proponen en este trabajo de tesis, todos los alumnos deben de tener acceso a la tecnología para facilitarles el aprendizaje bajo la guía de un profesor competente

A través de la tecnología puede potenciarse la implicación de los alumnos en las ideas matemáticas abstractas y en su dominio, también enriquece la gama y calidad de las investigaciones al

proveer medios para visualizar ideas matemáticas desde diversas perspectivas, se ayuda al aprendizaje mediante la retroalimentación p.e.; al arrastrar un punto en un entorno de geometría dinámica y observar como la imagen de la pantalla cambia, o al modificar las fórmulas en la hoja de cálculo y detectar como se modifican los valores dependientes. Así también permite centrar la atención cuando los alumnos discuten entre ellos o con su profesor sobre los objetos que aparecen en la pantalla y sobre los efectos de las posibles transformaciones dinámicas que el programa permite.

Muy en especial la tecnología ofrece posibilidades de adaptación de la enseñanza a las necesidades especiales del alumno en situación especial, las posibilidades para atraer a las matemáticas a los alumnos con discapacidades físicas aumentan radicalmente.

Shoenfeld (1998) identifica tres formas en que la tecnología puede ayudar a modificar las creencias de los estudiantes acerca de las matemáticas.

- “1.- La computadora permite vincular el trabajo empírico con lo formal en áreas como la geometría...
- 2.- Un aspecto esencial de las matemáticas es el analizar información estadística....
- 3.- Otra área importante en las matemáticas es la modelación de fenómenos”

Este trabajo toca un tema importante de la estadística el “análisis del modelo de regresión lineal”, que muchas veces no es abordado en el nivel superior con medios dinámicos en este sentido Santos (2000) cita a Shoenfeld (1992);

“...generalmente el estudiante separa una construcción de una demostración y pocas veces tiene la oportunidad de discutir la importancia y posibilidad de una construcción. Se recomienda que los alumnos participen en actividades en las que tengan que hacer construcciones y decidir cuándo una construcción no es posible”

La tecnología ofrece un potencial para ayudar al estudiante a profundizar en su aprendizaje. Por ejemplo, en el análisis de regresión lineal simple, la computadora ofrece un medio para que el estudiante explore, conjeture, analice y pruebe varias ideas relacionadas con ese contenido en particular, existe el medio para que desarrolle habilidades y estrategias que pueden posteriormente ser importantes en otros contextos.

La tecnología ayuda a que el estudiante no sea solo un espectador o receptor del conocimiento, sino que pasa a ser un ente activo vinculado directamente con el quehacer matemático, para que el uso de la tecnología pase a ser importante en la educación es necesario desarrollar una cultura tecnológica entre los maestros y estudiantes, en este sentido Santos (2000) señala;

“...El uso de la computadora ha influenciado notablemente la forma de desarrollar matemáticas. Por ejemplo, en la búsqueda de patrones o comportamiento de fenómenos, la computadora ha resultado

ser un gran instrumento que ayuda a representar y organizar lo que antes era difícil de sistematizar. Además, en algunos casos resulta relativamente simple variar los valores de una expresión algebraica para estudiar con detalle lo que pasa geoméricamente (representación gráfica)...”

Finalmente es importante mencionar que la tecnología debe ser vista como un gran apoyo en el aprendizaje de los estudiantes, pero en ningún momento como un sustituto del maestro, además el estudiante al usar el software debe entender los principios matemáticos que emerjan de determinada exploración y tener presente que muchas veces la tecnología puede dar una representación gráfica incompleta.

CAPITULO 3

METODOLOGÍA

En este capítulo se proponen las actividades a realizar para el aprendizaje del modelo de “regresión lineal” basados en un análisis didáctico, se hace un estudio desde tres perspectivas: el contenido, la enseñanza y el aprendizaje.

3.1 Actividades para la enseñanza del modelo de “regresión lineal”

Para seleccionar las actividades que lleven a los estudiantes al aprendizaje de los conceptos básicos del modelo de “regresión lineal” se propone comenzar con un tratamiento intuitivo que alcance a revisar conceptos de la estadística descriptiva y representación gráfica de datos con dos variables relacionadas. Después se pretende que a través de la representación gráfica y su análisis se llegue a una aproximación a la correlación lineal, para finalizar se incorporan los cálculos para hallar el coeficiente de correlación y a los coeficientes de la recta de regresión. También es deseable que los estudiantes lleguen a darle significado al coeficiente de correlación, a la recta de regresión y a la utilización de la recta ajustada para estimar valores de Y . El uso de la computadora se utiliza como una herramienta muy importante para el trabajo de la enseñanza se pretende que los estudiantes se concentren más en los significados de los resultados que en los procesos de cálculo.

Se presenta la propuesta didáctica sobre diferentes actividades para la enseñanza. La elección está sustentada en el trabajo de investigación realizado por A. Lavallo, E. Micheli, N. Rubio (2006). El enfoque presentado y los incisos que conforman las actividades están presentados en tres grupos, de acuerdo con los contenidos que involucran:

- 1) Datos con dos variables: Variables, descripción de las variables: unidad de medida, media y desvío estándar de cada una; Gráfico de dispersión: ubicación en el gráfico de la media de cada variable.
- 2) Relación entre variables-correlación: descripción de la relación entre dos variables a partir de un gráfico de dispersión, identificación de la relación lineal a partir de un gráfico de dispersión, caracterización del tipo y grado de relación entre variables a partir del gráfico de dispersión o de otra forma de representación, covariancia, cálculo del coeficiente de correlación lineal, interpretación del coeficiente de correlación lineal, correspondencia entre el valor del coeficiente de correlación propuesto a partir del análisis del gráfico y su valor calculado.
- 3) Regresión: cálculo de los coeficientes de la recta, gráfico de la recta de regresión, interpretación de los coeficientes de la recta, estimación de valores de Y cálculo del error de estimación y representación en el gráfico, interpretación del error como desvío.

A continuación, se presentan las actividades elegidas en cada sesión de clases previo un análisis didáctico, para este trabajo de tesis el grupo donde se aplican las actividades se conforma por estudiantes recién egresados del sistema Bachilleres en Chetumal.

A.1 Actividad 1

A.1.1 Análisis de contenido

En esta actividad se trata de acercarse al concepto de variable, identificar las unidades de medida e identificar los valores de las variables como pares ordenados, también se calcula la media y representa los valores de las variables gráficamente pasando de una presentación tabular a una presentación gráfica.

La relación entre los conceptos es, para graficar hay que tener en cuenta las unidades de medida de las variables y se pueden graficar los valores medios de cada variable.

A.1.2 Análisis de enseñanza

Esta actividad es una introducción al tema principal proponiendo trabajar con datos de dos variables presentados en una tabla, permitiendo observar los conocimientos previos de los estudiantes y que se puedan familiarizar con este tipo de datos, el nivel es elemental y se sugiere trabajar en computadora, en forma grupal y que el docente identifique errores a partir del resultado que presenten los estudiantes, en caso de ser necesario profundizar con actividades similares.

A.1.3 Análisis de aprendizaje

Al llevar a la práctica esta actividad se espera presenten los estudiantes errores como la interpretación de las variables en pares ordenados, cálculo de las medias o al elegir la escala en el gráfico, en esta actividad también se espera conocer la profundidad de conocimientos previos en cada estudiante respecto a la estadística.

A.1.4 Enunciado

Jensen Tire & Auto está por decidir si firma un contrato de mantenimiento para su nueva máquina de alineamiento y balanceo de neumáticos. Los gerentes piensan que los gastos de mantenimiento deben estar relacionados con el uso y recolectan los datos siguientes sobre el promedio de uso semanal (horas) por cada año y gastos anuales de mantenimiento (en cientos de dólares).

Tabla 3.- datos sobre el promedio de uso semanal (horas) por cada año y gastos anuales de mantenimiento (en cientos de dólares).

<i>Uso semanal</i>	<i>Gasto Anual de mantenimiento</i>
13	17
10	22
20	30
28	37
32	47
17	30.5
24	32.5
31	39
40	51.5
38	40

- a) *¿Cuáles son las variables de interés para los gerentes?*
- b) *¿Qué unidad de medida se utiliza para calcular cada variable?*
- c) *¿Qué significa el par ordenado (28; 37)?*
- d) *¿Cuál es el tiempo promedio de uso de la máquina?*
- e) *¿Cuál es el gasto anual medio de mantenimiento?*
- f) *Grafica los puntos de la tabla e indica en el mismo los valores hallados en los incisos c), d) y e).*

A.2 Actividad 2

A2.1 Análisis de contenido

En esta actividad se refuerza el concepto del promedio, así como elaborar el gráfico de dispersión donde se identifique si hay relación (o no) entre las variables, que tipo de relación (lineal, no lineal, directa, inversa) y el grado de relación lineal.

Observando el gráfico de dispersión el estudiante puede identificar si existe un patrón de relación lineal entre las variables, también con esta actividad se pretende que el estudiante tenga noción del coeficiente de correlación lineal como medida sobre el grado de intensidad de la relación lineal existente entre las variables, a medida que los puntos están más alineados, la correlación entre las variables es más fuerte., Esta actividad permite transitar de la forma tabular a la gráfica y de la gráfica a la descripción a través de la interpretación sobre la relación dada.

A.2.2 Análisis de enseñanza

Al desarrollar esta actividad en el aula pueden surgir dificultades en la elaboración de los gráficos de dispersión, donde tenga que intervenir el docente, se retoman contenidos de estadística descriptiva, al asignar un valor de correlación, se avanza con respecto a la actividad anterior, es muy importante que el estudiante conozca el recorrido del coeficiente y su interpretación en relación con su magnitud y signo.

A.2.3 Análisis de aprendizaje

Se intenta que el estudiante realice una buena lectura de los datos, es decir, que grafiquen respetando las escalas e identifiquen las variables correspondientes a cada eje. La lectura dentro de los datos involucra el cálculo y representación en el gráfico de la media de cada variable y la determinación del tipo y grado de relación, observando el gráfico de dispersión, se espera que el estudiante supere posibles errores en la aproximación del coeficiente de correlación en especial en gráficos de dispersión poco densos.

A.2.4 Enunciado

Las donaciones de los ex alumnos son una importante fuente de ingresos para las universidades. Si los administradores pudieran determinar los factores que influyen sobre el aumento de la cantidad de alumnos que hacen donaciones, podrían poner en marcha políticas que llevarían a ganancias mayores. Las investigaciones indican que los estudiantes más satisfechos con la relación con sus profesores tienen más probabilidad de titularse, lo que a su vez lleva al aumento de la cantidad de alumnos que hagan donaciones. En la tabla siguiente se muestran datos de 12 universidades de Estados Unidos (American's Best Collage, edición del año 2000). La columna titulada "tasa de titulados" da el porcentaje de alumnos titulados, de los inicialmente inscritos. La columna que tiene como título "% de grupos con menos de 20" da muestra del porcentaje de grupos con menos de 20 alumnos. La columna que tiene como título "Tasa de estudiantes/facultad" da la cantidad total de estudiantes inscritos, dividida entre la cantidad de facultades. Por último, la columna que tiene como título "Tasa de alumnos que donan" da el porcentaje de alumnos que han hecho alguna donación a la universidad.

Tabla 4.- En la tabla siguiente se muestran datos de 12 universidades de Estados Unidos (American's Best Collage, edición del año 2000)

Universidad	Tasa de titulados	% de grupos con menos de 20	Tasa de estudiantes/facultad	Tasa de alumnos que donan
Boston College	85	39	13	25
Brandeis University	79	68	8	33
Brown University	93	60	8	40
California Institute of Technology	85	65	3	46
Carnegie Mellon University	75	67	10	28
Case Western Reserve University	72	52	8	31
College of William and Mary	89	45	12	27
Columbia University	90	69	7	31
Cornell University	91	72	13	35
Dartmouth College	94	61	10	53
Duke University	92	68	8	45
Emory University	84	65	7	37

a) *Elabora tres gráficos de dispersión donde en el eje de las ordenadas de cada uno de ellos figure la tasa de alumnos que donan y en el eje de las abscisas i) tasa de titulados, ii) % de grupos con menos de 20 y iii) tasa de estudiantes/facultad.*

b) *¿Qué puedes decir, en cada caso, sobre el tipo de relación entre las variables consideradas?*

c) *¿Cuál es el promedio de la tasa de titulados? ¿Cuál el de % de grupos con menos de 20? ¿Cuál el de tasa de estudiantes/facultad? Indica en los gráficos.*

d) *¿Cuál es la tasa promedio de los alumnos que donan en estas universidades?*

e) *Indica cuál de los siguientes valores te parece más próximo al coeficiente de correlación lineal del inciso i): -0,98; 0,54; 0,98; 0,73.*

f) *Determina un valor aproximado de r para los incisos ii) y iii).*

A.3 Actividad 3

A3.1 Análisis de contenido

Esta actividad trata de reforzar el conocimiento de identificar cuando existe relación entre dos variables a partir de la observación del gráfico de dispersión e identificar si la relación es lineal o no, así como reforzar el concepto de correlación (directa, inversa), insistir en que la correlación “ r ” es una medida sobre el grado de intensidad de la relación lineal existente entre las variables, a medida que los puntos están más alineados, la correlación entre las variables es más fuerte, si r es aproximadamente igual a 0 y los datos tienen un patrón lineal, entonces los puntos caen en una recta horizontal.

A3.2 Análisis de enseñanza

Dado que la actividad no exige cálculos sino la interpretación de gráficas y la determinación aproximada del coeficiente de correlación se puede considerar una actividad intuitiva y se considera una buena opción realizarla en grupo para proponer una discusión entre los estudiantes sobre sus respuestas. Aquí el docente debe guiar la discusión y evaluar las aproximaciones de “ r ”.

A3.3 Análisis de aprendizaje

Los conocimientos previos que son necesarios para esta actividad se refieren sólo a la relación entre variables. Se intenta colaborar en la superación de errores relacionados con el ejercicio de proporcionar un valor aproximado del coeficiente de correlación, con lo que se avanza respecto a la tarea anterior. Para identificar los distintos tipos de relación y poder proponer un valor aproximado del coeficiente de correlación es necesaria una lectura dentro de los datos.

A3.4 Enunciado

Observa los siguientes gráficos y responde:

a) ¿En cuáles de ellos puedes decir que existe relación entre las variables?

b) ¿En cuáles existe relación lineal?

c) En los gráficos que muestran relación lineal entre las variables, analiza el tipo y grado de esa relación y, en cada caso, determina un valor aproximado del coeficiente de correlación.

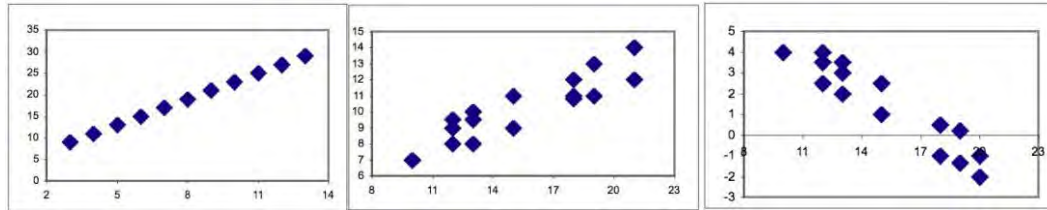


Fig., a

Fig., b

Fig., c

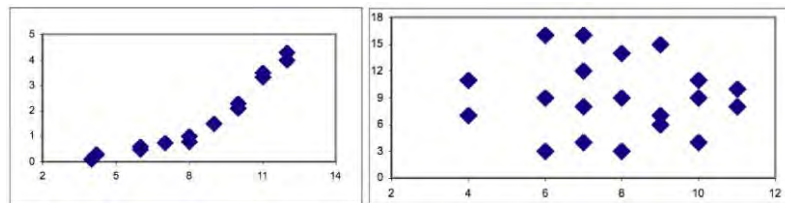


Fig. d

Fig. e

Figura 3.- gráficos de dispersión para la actividad tres

A.4 Actividad 4

A4.1 Análisis de contenido

Con esta actividad se pretende reforzar el concepto de variable, como es identificar las variables en estudio y distinguir entre una variable respuesta y una variable explicativa, interpretando los gráficos de dispersión en juego.

A4.2 Análisis de enseñanza

Esta actividad exige un nivel de comprensión mayor, ya que los enunciados pueden generar dificultades a los alumnos, resulta interesante el tipo de ejercicio en el que se parte de una representación verbal. Es recomendable trabajar en grupos para posibilitar la discusión sobre los enunciados.

A4.3 Análisis de aprendizaje

Es importante que los estudiantes se familiaricen con el significado de indicadores estadísticos, es posible que surjan errores asociados con la falta de interpretación de dichos indicadores, para interpretar y resolver los enunciados se requiere de una comprensión tanto de los gráficos que involucran la lectura dentro de los datos, como del tipo de relación enunciado verbalmente. Los conocimientos previos necesarios no son diferentes a los requeridos en las actividades anteriores.

A4.4 Enunciado

A continuación, se presentan tres afirmaciones referidas a las conclusiones de un estudio acerca de las tasas de nacimiento, suicidio, crecimiento económico y productividad, junto con tres gráficos de dispersión.

Afirmación 1: En países con un desarrollo tecnológico alto, como Japón, Estados Unidos, Alemania, Inglaterra, Francia, Italia y Canadá se tienen bajas tasas de nacimiento (TN) asociadas con altas tasas de suicidio (TS).

Afirmación 2: Algunos economistas afirman que, independientemente de los países que se estudian, a altas tasas de crecimiento (TC) se asocian altas tasas de productividad (TP).

Afirmación 3: Tanto economistas como demógrafos afirman que las tasas de suicidio (TS) no parecen estar correlacionadas con las tasas de productividad (TP). Desafortunadamente, en los gráficos no se colocaron los rótulos de referencia de los ejes. Asocia a cada gráfica una afirmación, completa con el nombre de los ejes y proporciona en cada caso un valor del coeficiente de correlación.



Figura 4.- gráficos de dispersión para la actividad cuatro

A.5 Actividad 5

A5.1 Análisis de contenido

Esta actividad fue propuesta por Sánchez Cobo et al. (2000), tiene como propósito que el estudiante traduzca la información brindada por una fórmula a un diagrama, entendiendo que el coeficiente de correlación es una medida sobre el grado de intensidad de la relación lineal existente entre las variables, a medida que los puntos están más alineados, la correlación entre las variables es más fuerte, si r es aproximadamente igual a 0 y los datos tienen un patrón lineal, entonces los puntos caen en una recta horizontal.

A5.2 Análisis de enseñanza

La actividad exige un tratamiento previo acerca del coeficiente de correlación, como los mencionados en las actividades anteriores, como el proceso inverso que, dado un diagrama de dispersión, sugerir un valor razonable del coeficiente. El docente puede observar a partir de los resultados si los alumnos relacionan valores cercanos a cero, pero negativos (inciso c), con el concepto de independencia de variables, y si elaboran diagramas acertados para valores intermedios del coeficiente de correlación.

A5.3 Análisis de aprendizaje

Los diferentes incisos involucran distintas intensidades de la correlación. En particular, el b) expresa una relación funcional y el c) independencia. De esta forma, se puede observar si a mayor intensidad mejora la interpretación del tipo de relación, y si se interpretan correctamente los valores del coeficiente de correlación que no son valores extremos.

A5.4 Enunciado

Dados los siguientes valores del coeficiente de correlación entre dos variables x , y , dibuja un diagrama de dispersión, con 10 puntos, que se adapte razonablemente a ellos.

a) $r = 0,25$

d) $r = 0,7$

b) $r = -1$

e) $r = -0,4$

c) $r = -0,01$

A.6 Actividad 6

A6.1 Análisis de contenido

En esta actividad de cierre se pretende transitar por los conceptos de la regresión entre ellos, verificar a partir del gráfico la conjetura de relación lineal entre las variables, calcular los coeficientes de la recta, graficar la recta, interpretar los coeficientes de la recta, calcular valores estimados de “ y ”, indicar gráficamente el error de estimación, interpretar el error como desvío, calcular el coeficiente de correlación, interpretar valores del coeficiente de correlación.

Las interrelaciones entre los conceptos aluden a que en la regresión lineal las variables están relacionadas mediante la ecuación de la recta de regresión, la recta de regresión estimada es la que se ajusta lo mejor posible a la nube de puntos del gráfico de dispersión (en el sentido de los mínimos cuadrados). el valor estimado de Y a partir de los valores de X se obtiene con

$$Y_i = \beta_0 + \beta_1 X_i, \text{ la pendiente de la recta indica la variación que se produce en } Y \text{ cuando}$$

X varía una unidad; esto debe interpretarse de acuerdo con las variables y el rango de valores estudiados, la ordenada al origen de la recta indica cuánto vale Y cuando $X=0$, lo cual debe interpretarse de acuerdo con las variables y el rango de valores estudiados, el coeficiente de correlación auxilia a determinar si hay relación lineal entre las variables.

En esta actividad, se propone pasar de la forma de representación tabular a la gráfica, al igual que realizar un tratamiento general de los conceptos involucrados. La actividad comienza a nivel intuitivo, pero permite comprobar las aproximaciones a través del cálculo de los coeficientes correspondientes.

A6.2 Análisis de enseñanza

Esta actividad, que reúne todos los conceptos y procedimientos de las anteriores, se propone que se desarrolle en forma grupal, lo cual posibilita la discusión y el intercambio de ideas. Es posible plantear actividades similares en las que se trabaje con los datos recabados por los propios alumnos, provenientes de variables de su interés.

A6.3 Análisis de aprendizaje

Ahora se propone agregar un nivel de dificultad mayor, pero que al mismo tiempo sea utilizado como actividad de cierre. Basándose en el cálculo, el alumno puede realizar una predicción del valor de una variable en función de la otra. Aquí, el nivel de lectura que se pretende es aquel que va más allá de los datos, en un estadio de conocimiento mucho más profundo que el de las lecturas anteriores.

Tomando el enunciado de la actividad uno realiza lo siguiente

- a) *Elaborar un gráfico de dispersión para los datos representados en la tabla del enunciado.*
- b) *Observa el gráfico. Decide si es razonable suponer que existe relación lineal entre las variables; si es así, obtenga la ecuación de regresión estimada que relaciona gastos anuales de mantenimiento con el uso semanal e interpreta los coeficientes. Grafica la recta y estima una aproximación del coeficiente de correlación.*
- c) *Calcula e interpreta el coeficiente de correlación. ¿Estuvo acertada la aproximación que hiciste en el inciso anterior?*
- d) *Jensen piensa que usará la nueva máquina 30 horas a la semana. En ese caso, ¿Qué monto sería el gasto anual de mantenimiento de la empresa?*
- e) *Si el precio del contrato de mantenimiento es \$3000 anuales, ¿recomendaría firmar el contrato de mantenimiento?, ¿Por qué sí o por qué no?*
- f) *Si quisieras estimar el gasto anual de mantenimiento para un uso semanal de la máquina de 5 horas, ¿puedes usar la recta que calculaste en el inciso b? ¿Por qué?*
- g) *¿Qué error se comete en la estimación realizada en el inciso? Indícalo en el gráfico.*

A.7 Actividad 7

Con esta actividad se propone evaluar el nivel de conocimiento alcanzado por los estudiantes en las sesiones anteriores, la actividad se resolverá de manera individual, para después cerrar con una discusión en grupo respecto a los conceptos que transitaron en el aula.

A7.1 Enunciado

Decide si las siguientes sentencias son verdaderas o falsas y justifica brevemente:

- a) Un coeficiente de correlación igual a 1 entre “x, y” significa que “x” está relacionada con “y”, pero un coeficiente de correlación -1 significa que “x” no está relacionada con “y”.*
- b) Si la pendiente de la recta es positiva, el coeficiente de correlación es positivo.*
- c) El signo del coeficiente de correlación depende del signo de la covarianza.*
- d) Un coeficiente de correlación igual a cero indica una no dependencia lineal entre las variables.*
- e) Si existe una correlación alta, se concluye que “x” es causa de “y”.*
- f) Si el coeficiente de correlación es cercano a 1, “x” es directamente proporcional a “y”.*

CAPITULO 4

RESULTADO

En esta sección se describe principalmente los resultados obtenidos en el desarrollo del conocimiento del tema en los estudiantes durante la implementación de las actividades de aprendizaje para regresión lineal propuesta en la secuencia didáctica, se analizan los datos obtenidos, procedimientos y dificultades en los estudiantes, así como el papel del profesor como guía facilitador para que el ambiente en la clase se acerque a lo planeado y la respuesta a las dificultades contribuyan en el desarrollo del conocimiento de los estudiantes.

4.1 Análisis de las actividades

Las actividades de la secuencia didáctica se aplicaron en estudiantes recién egresados del nivel bachillerato, como herramienta de apoyo se les solicitó a los estudiantes llevar su computadora portátil en cada sesión, prácticamente los conocimientos de estadística que tienen son los desarrollados en el nivel básico (secundaria), aunque se pudo identificar que, a sentimiento de ellos, no sabían nada de estadística. La implementación de las actividades fue en cuatro sesiones, la primera sesión fue de una hora y las siguientes de dos horas, en general las sesiones se desarrollaron en un ambiente colaborativo entre estudiantes y docente, La secuencia didáctica completa de las actividades que se desarrollaron en el aula se describen en el anexo “A”, la implementación de las actividades en este apartado se describen intentando captar el desarrollo del conocimiento del tema, las dificultades encontradas, el papel del docente como guía y el ambiente de colaboración entre estudiantes, en cuanto a la tecnología de cómputo, se utilizó la hoja de Excel para tablas y gráficas y el programa estadístico “R” Commander en un nivel muy básico para tablas y gráficos de dispersión aunque los estudiantes conocían muy poco de gráficas en Excel y nada de “R” , no fue un obstáculo para el desarrollo de la secuencia, todo lo contrario aprendieron a sacarle provecho muy pronto, el desarrollo de cada sesión se describe en seguida.

SESIÓN 1

ACTIVIDAD 1

Esta primera actividad sirvió como introducción al tema principal, se trabajó a partir de un enunciado(ver anexo A), con datos de dos variables presentados en una tabla, esta actividad permitió observar los conocimientos previos de los estudiantes y al mismo tiempo se pudieron familiarizar con este tipo de datos, el nivel fue muy elemental lo que propició confianza en los estudiantes, se trabajó en computadora de manera individual en un principio y al cierre en forma grupal, donde el docente pudo identificar dificultades a partir de las respuestas que presentaron los estudiantes, después de aclarar dudas e inexactitudes el docente profundizó en el tema del promedió. La descripción del desarrollo de la clase se presenta con las respuestas escritas de cada reactivo y como se desarrolló su respectivo análisis en forma grupal al cierre de la clase. Para iniciar la actividad se entregó a los estudiantes una hoja conteniendo el enunciado, los datos y cinco preguntas, se les indicó que resolvieran las preguntas y actividades solicitadas para después discutir las en grupo (ver anexo A) las preguntas y respuestas con las que se trabajaron son las siguientes:

La primera pregunta fue:

a) ¿Cuáles son las variables de interés para los gerentes?

Las respuestas escritas fueron semejantes a lo siguiente:

Estudiante 1; “las variables son el uso y el gasto”

Estudiante 2; “uso semanal, gasto anual de mantenimiento”

La segunda pregunta fue:

b) ¿Qué unidad de medida se utiliza para calcular cada variable?

Las respuestas en este caso fueron similares a:

Estudiante 1; “horas y dólares”

Estudiante 2; “uso semanal: horas/semana, gasto anual de mantenimiento: ciento de dólares/año”

En las respuestas como la del estudiante 1 se puede observar las dificultades que manifiestan con la interpretación de las variables del enunciado, tiene que ver con el carácter lingüístico, porque implica una interpretación incorrecta del lenguaje usado al describir las variables en la respuesta de la primera pregunta y las unidades en la respuesta a la segunda pregunta. Las repuestas del estudiante 2 están mejor descritas, aunque nadie mencionó en su respuesta a la pregunta 1 la palabra clave “promedio”, esto se discutió en la actividad final de la clase.

En la actividad final de esta sesión se intercambiaron las respuestas de manera verbal entre los estudiantes, el docente indujo a una discusión tomando como referencia las respuestas anteriormente mencionadas preguntando si ambas respuestas eran correctas (en cada pregunta). Después del intercambio de opiniones, como puesta en común, estuvieron de acuerdo en que las respuestas deben ser más descriptivas (como las del Estudiante 2), tanto para las variables como para las unidades de medida.

Las siguientes preguntas y sus respuestas representativas fueron:

c) ¿Qué significa el par ordenado (28; 37)?

Estudiante 1; “si el mantenimiento se usa 28 horas a la semana genera un gasto anual de 3700 dólares”

Estudiante 2; “significa, usando semanalmente 28 horas el automóvil hay un gasto anual de 3700 dólares”

d) ¿Cuál es el tiempo promedio de uso de la máquina?

Estudiante 1: “25.3 horas”

Estudiante 2: “uso promedio= $(13+10+20+28+32+17+24+31+40+38) /10= 253/10= 25.3$ el tiempo promedio de uso de la máquina es de 25.3 horas semanalmente”

e) ¿Cuál es el gasto anual medio de mantenimiento?

Estudiante 1: “34.65 dólares”

Estudiante2: “gastomed= $(17+22+30+37+47+30.5+32.5+30+51.5+40) /10=3465/10=34.65$ el gasto anual medio de mantenimiento es de 34.65 dólares”

Se puede observar en las respuestas del inciso c) que si correlacionaron intuitivamente el par de variables pero estaban un poco fuera del contexto del enunciado, en los siguientes dos incisos el estudiante 1 utilizó su computadora para encontrar el promedio mientras que el estudiante 2 utilizó la calculadora para encontrar el promedio y fue más descriptivo en su respuesta, las respuestas muestran que si tienen conocimientos previos en cuanto a estadística en especial en promedios y se comienza a involucrar en el estudiante la noción de correlación.

En el cierre de la sesión se indujo a la discusión de que significa el valor promedio, su utilidad, si creen que pueda servir como parámetro para predecir un valor futuro, se observa las ideas de los estudiantes, con intención no se llega a una respuesta concreta sugiriendo a los estudiantes que en las

siguientes sesiones lo descubriremos y por otra parte se intercambia opiniones respecto al significado del enunciado (problema principal).

La última actividad del enunciado fue lo siguiente:

f) Grafica los puntos de la tabla e indica en el mismo los valores hallados en los incisos c), d) y e)

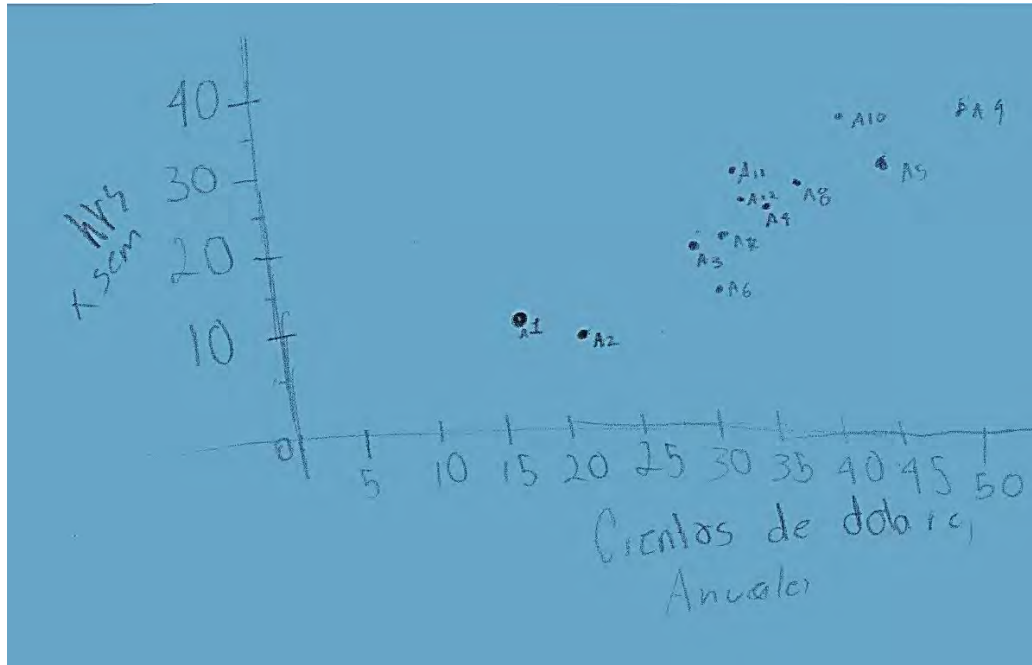


Figura 5.- Gráfica de puntos elaborada por estudiante 1.

En esta actividad a pregunta expresa de los estudiantes, se les indicó que el gráfico lo pueden hacer como ellos decidieran, en la fig. “x1” se observa el gráfico desarrollado por un estudiante en lápiz y papel, que, aunque numéricamente es correcto incluyendo los puntos promedio, no hay la intuición de que “el gasto anual promedio” puede depender del “uso semanal de la máquina” observando los ejes del gráfico.

En el cierre de la sesión se intercambiaron entre los estudiantes los gráficos en papel, el docente propone al grupo describir el gráfico, obteniendo respuestas interesantes entre ellas “son puros puntos al azar”, el docente insiste creen que tengan alguna tendencia, obteniendo respuestas como “van hacia arriba (los puntos)” o “no veo nada”, después se propone hacer el gráfico en grupo en la hoja de cálculo de Excel con las indicaciones del inciso f) resultando el siguiente gráfico:

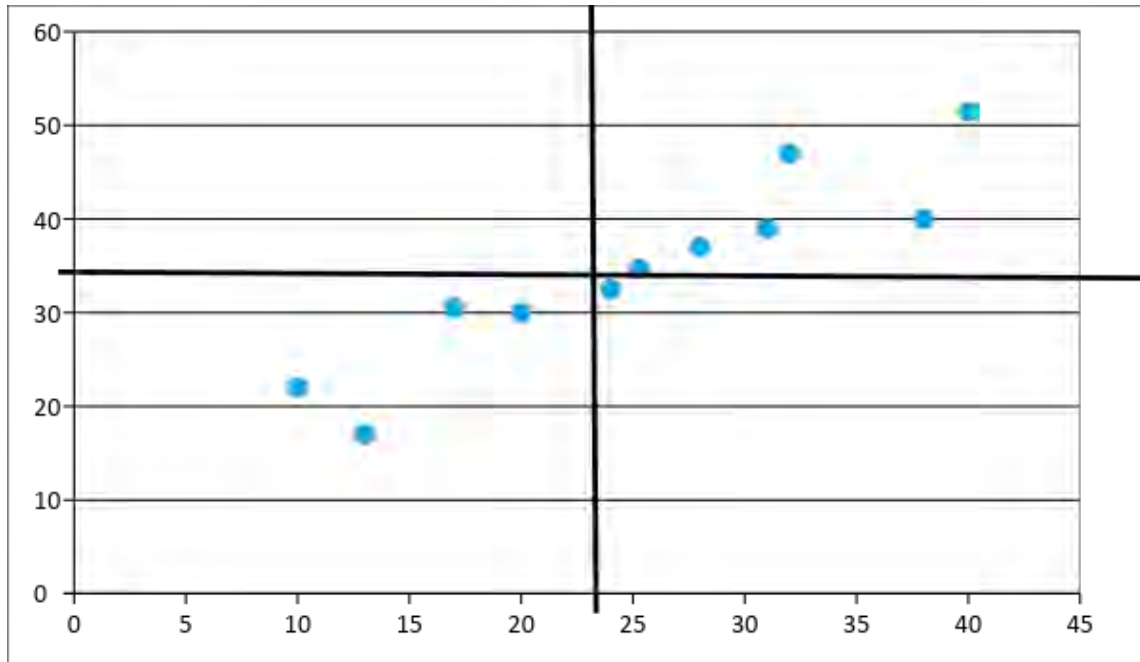


Figura 6.- gráfico de dispersión elaborado en grupo en la hoja de cálculo de Excel con las indicaciones del inciso f)

A partir de esta figura el docente expuso los aspectos del gráfico de puntos mencionando que también se le conoce como “gráfico de dispersión” y su comprensión es indispensable para entender el objetivo principal, también se expuso sin profundizar en el tema como observando en que cuadrantes caen la mayoría de los puntos se dice que tienen tendencia positiva o tendencia negativa, para finalizar la sesión se les presentó el programa estadístico “R Commander” y se les dio las instrucciones para elaborar el mismo gráfico en ese programa como tarea de extra clase.

SESIÓN 2

ACTIVIDAD 2

La segunda actividad sirvió para introducir el tema de correlación, se trabajó a partir de un enunciado (ver anexo A), con datos presentados en una tabla (Tabla 4), esta actividad permitió observar el desarrollo en la interpretación y descripción de los datos presentados, se acercan a la identificación de las variables que deben corresponder en cada eje (en el gráfico de dispersión), se acercan a identificar la importancia de usar escalas adecuadas. La lectura de los datos involucró también el cálculo y representación en el gráfico de la media de cada variable, se trabajó en el programa estadístico “R” y en la hoja de cálculo Excel e intuitivamente los estudiantes se acercaron en la determinación del tipo y grado de relación observando el gráfico de dispersión.

La primera actividad se trabajó de manera individual y después se reflexionó en grupo, fue de acuerdo a las siguientes instrucciones:

a) *Elabora tres gráficos de dispersión donde en el eje de las ordenadas de cada uno de ellos figure la tasa de alumnos que donan y en el eje de las abscisas*

i) tasa de titulados, ii) % de grupos con menos de 20 y iii) tasa de estudiantes/facultad.

b) *¿Qué puedes decir, en cada caso, sobre el tipo de relación entre las variables consideradas?*

c) *¿Cuál es el promedio de la tasa de titulados? ¿Cuál el de % de grupos con menos de 20? ¿Cuál el de tasa de estudiantes/facultad? Indica en los gráficos.*

d) *¿Cuál es la tasa promedio de los alumnos que donan en estas universidades?*

Para iniciar la reflexión en grupo se leyeron las respuestas del inciso b) encontrando argumentos como: “... de la tasa de titulados, solo un porcentaje de estos dona a la universidad. Y así en todos los casos...”

Para los gráficos de la actividad individual hubo quien trabajó en Excel y quien trabajó en “R”

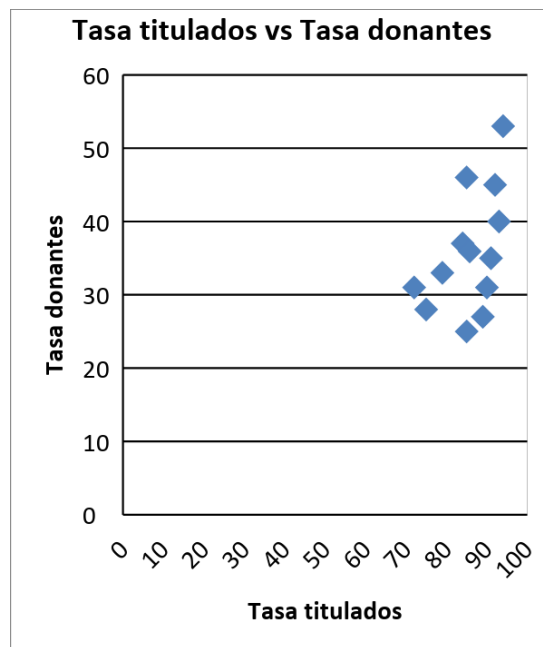


Figura 6a.- gráfico elaborado individualmente por estudiante 1

En la figura 6a se presenta el gráfico en Excel del inciso i), la diferencia con el programa “R” es la “escala”, en este caso se nota claramente la tendencia y correlación de los datos sin embargo no da la

oportunidad de *analizar* claramente la dispersión con respecto a los cuadrantes trazados desde el punto que se ubica en la media de cada variable

La figura 6b representa el gráfico del inciso *ii)* el cual dio la oportunidad al docente de reflexionar sobre la tendencia (directa o inversa) de los datos e intuitivamente sobre el grado de correlación llamado “r”

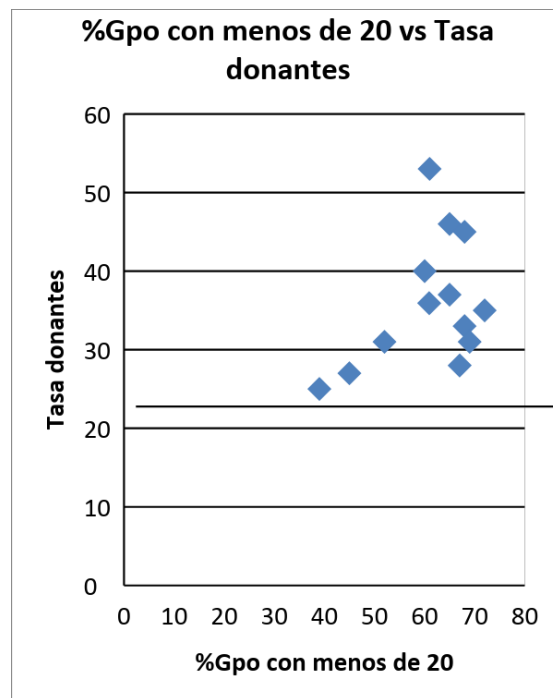


Figura 6b.- gráfico elaborado individualmente por estudiante 1

La figura 6c también dio la oportunidad de reflexionar ahora en grupo sobre la tendencia de los datos y el grado de correlación (se insiste que se conoce como “r”)

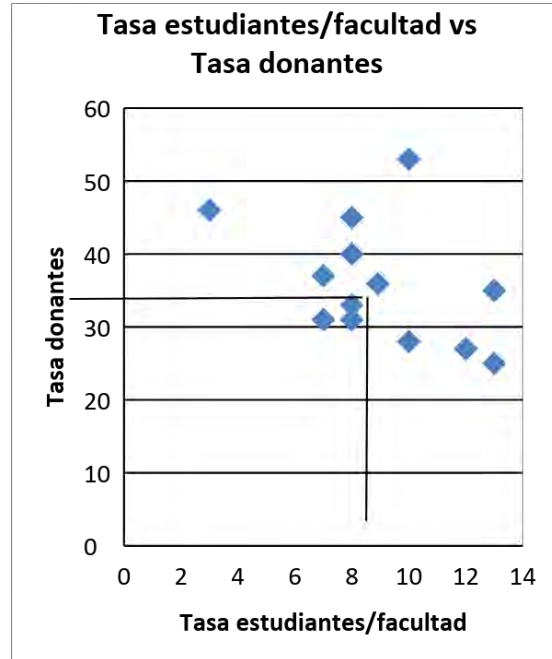


Figura 6c.- gráfico elaborado individualmente por estudiante 1

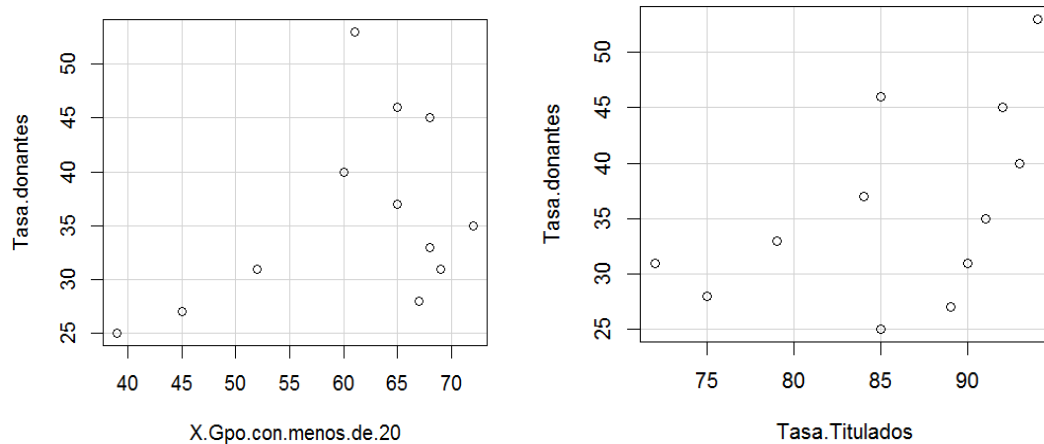


Figura 6d.- gráfico elaborado individualmente por estudiante 2

Los gráficos que se obtuvieron con el programa “R” son los que se indican en las figuras 6d y 6e, como se puede observar la escala de cada gráfico es más adecuada y se puede visualizar mejor la

dispersión de los datos e intuitivamente el grado de correlación, en este caso el docente propone trazar una línea imaginaria en el gráfico que pueda representar a ese conjunto de puntos y observar, mientras más se acerquen la dispersión de puntos a la línea imaginaria se dice que la correlación se acerca a “1” (uno) y mientras los puntos estén más dispersos con respecto a la línea imaginaria se dice que la correlación se acerca a “0” (cero), en estos casos la figura de arriba es tendencia directa y la figura de abajo es tendencia inversa, la participación de los estudiantes fue muy importante para acercarse al concepto de correlación lineal entre dos variables, el ambiente en la actividad en grupo fue con mucha participación, dejando por momentos que entre los estudiantes opinen sobre sus conclusiones y el docente procurando que la discusión no se aleje del tema principal, para concluir la sesión se les indicó a los estudiantes que respondieran los incisos e) y d).

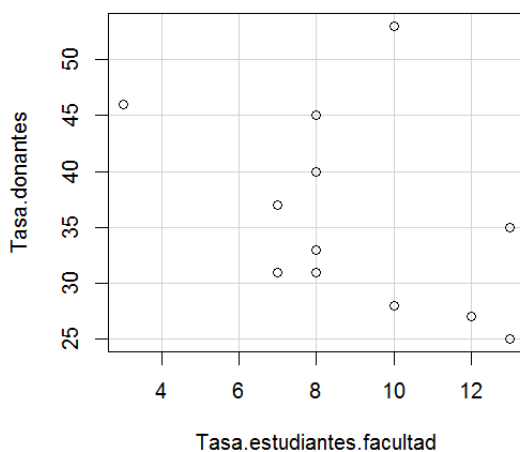


Figura 6e.- gráfico elaborado individualmente por estudiante 2

e) Indica cuál de los siguientes valores te parece más próximo al coeficiente de correlación lineal del inciso i): -0,98; 0,54; 0,98; 0,73.

Obteniendo respuestas como:

Estudiante 1: -0.98

Estudiante 2: 0.73

Es interesante observar en las respuestas, el estudiante 1 respondió en base a su diagrama realizado en Excel pues dado la escala que utiliza el programa aparenta estar más agrupados los puntos que lo que realmente lo están, utilizando una escala adecuada como en el caso del estudiante 2 que basó su respuesta en base a su gráfico realizado en el programa “R” se observa que percibe una dispersión de puntos moderada (que se acerca más a lo real) para el caso de alcanzar el objetivo de la sesión se observa un avance significativo.

f) Determina un valor aproximado de r para los incisos ii) y iii).

Las respuestas fueron:

Estudiante 1: ii) (entre) 0.50 y 0.20 iii) (entre) -0.50 y -0.70

Estudiante 2: ii) 0.54 iii) -0.00

En las respuestas se observa cómo se van acercando intuitivamente a la noción del coeficiente de correlación lineal, identifican cuando su tendencia (pendiente) es positiva o negativa, en la respuesta del inciso iii) del estudiante 2 percibe una gran dispersión y lo toma como coeficiente “0” que aunque no es correcto fue una respuesta de acuerdo a lo que transitó en la clase, esto se aprovecha en la siguiente clase para comentar el caso especial cuando el coeficiente de correlación es “0” “si r es aproximadamente igual a 0 y los datos tienen un patrón lineal, entonces los puntos caen en una recta horizontal”

y también se aprovecha para abundar acerca del coeficiente de correlación lineal.

SESIÓN 3

ACTIVIDAD 3

Esta actividad se puede considerar intuitiva, ya que requirió únicamente de la interpretación de gráficos, se realizó en principio individual y después se propuso una discusión en grupo, el docente sirvió de guía en la discusión y evaluación en las aproximaciones de “r”, se intentó reforzar el conocimiento de identificar cuando existe relación entre dos variables a partir de la observación del gráfico de dispersión e identificar si la relación es lineal o no, así como reforzar el concepto de correlación (directa, inversa) e insistir en que la correlación lineal “r” es una medida sobre el grado de intensidad de la relación lineal existente entre las variables, “a medida que los puntos están más alineados, la correlación entre las variables es más fuerte”. La actividad (ver anexo A) consistió en:

Observa los siguientes gráficos y responde:

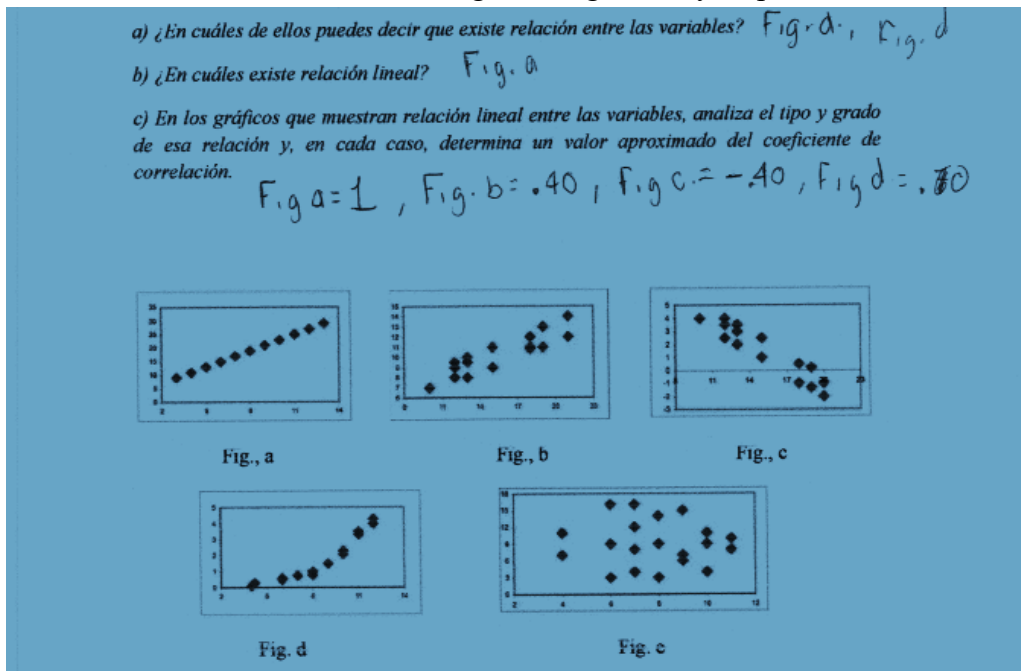


Figura 7.- Actividad 3 elaborada por estudiante 1.

ACTIVIDAD 4

En esta actividad se trabajó en un principio de manera individual y después se propuso la discusión en grupo, el trabajo individual exigió un nivel de comprensión mayor con respecto a las actividades anteriores, sirvió para reforzar el concepto de variable, como es identificar las variables en estudio y distinguir entre una variable respuesta y una variable explicativa interpretando los gráficos de dispersión, aunque los estudiantes tuvieron dificultades en la interpretación de los enunciados, resulta interesante partir de una representación verbal para distinguir su gráfico. La actividad (ver anexo A) consistió en:

A continuación, se presentan tres afirmaciones referidas a las conclusiones de un estudio acerca de las tasas de nacimiento, suicidio, crecimiento económico y productividad, junto con tres gráficos de dispersión.

Afirmación 1: En países con un desarrollo tecnológico alto, como Japón, Estados Unidos, Alemania, Inglaterra, Francia, Italia y Canadá se tienen bajas tasas de nacimiento (TN) asociadas con altas tasas de suicidio (TS).

Afirmación 2: Algunos economistas afirman que, independientemente de los países que se estudian, a altas tasas de crecimiento (TC) se asocian altas tasas de productividad (TP).

Afirmación 3: Tanto economistas como demógrafos afirman que las tasas de suicidio (TS) no parecen estar correlacionadas con las tasas de productividad (TP). Desafortunadamente, en los gráficos no se colocaron los rótulos de referencia de los ejes. Asocia a cada gráfica una afirmación, completa con el nombre de los ejes y proporciona en cada caso un valor del coeficiente de correlación.

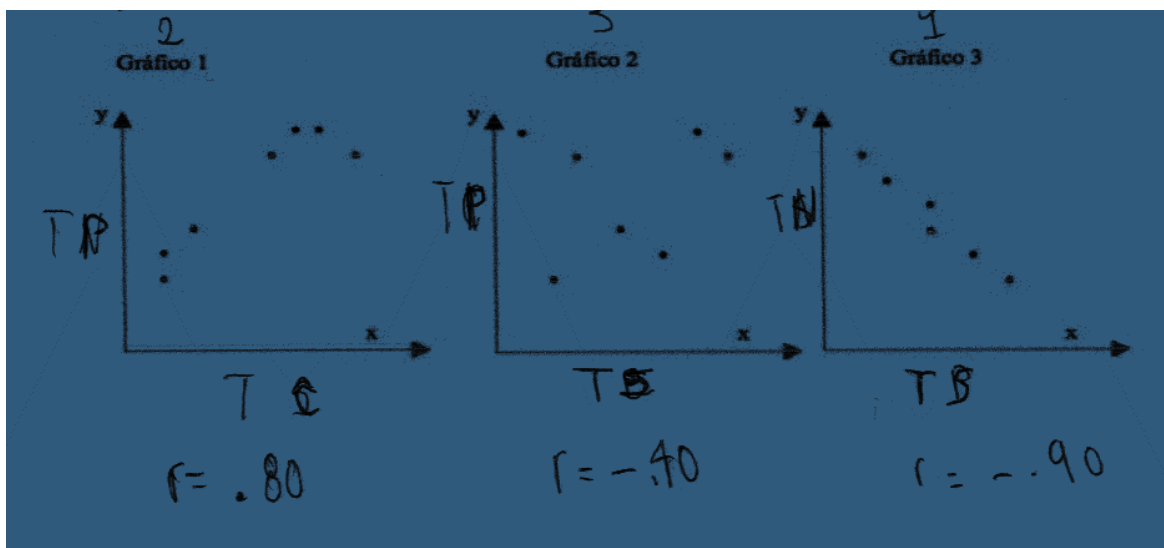


Figura 8.- Actividad 4 elaborada por estudiante 1.

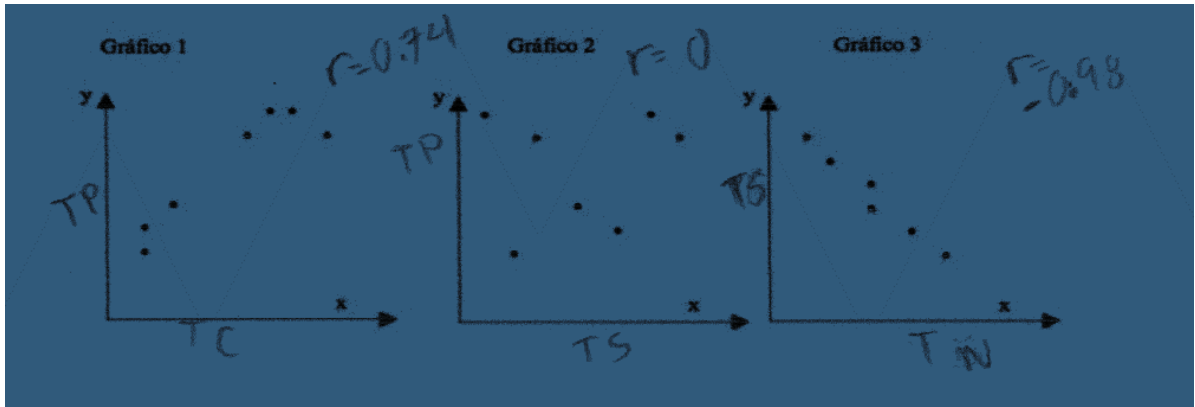


Figura 9.- Actividad 4 elaborada por estudiante 2.

ACTIVIDAD 5

Como cierre de la sesión para evaluar el desarrollo del conocimiento en el tema e identificar dificultades para su aprendizaje en los estudiantes se les presentó una actividad que exige un tratamiento previo acerca del coeficiente de correlación, como los mencionados en las actividades anteriores, como el proceso inverso que, dado un diagrama de dispersión, sugerir un valor razonable del coeficiente. La actividad (ver anexo A) consistió en:

Dados los siguientes valores del coeficiente de correlación entre dos variables x , y , dibuja un diagrama de dispersión, con 10 puntos, que se adapte razonablemente a ellos.

- | | |
|----------------|---------------|
| a) $r = 0,25$ | d) $r = 0,7$ |
| b) $r = -1$ | e) $r = -0,4$ |
| c) $r = -0,01$ | |



Figura 10.- Actividad 5 elaborada por estudiante 1.

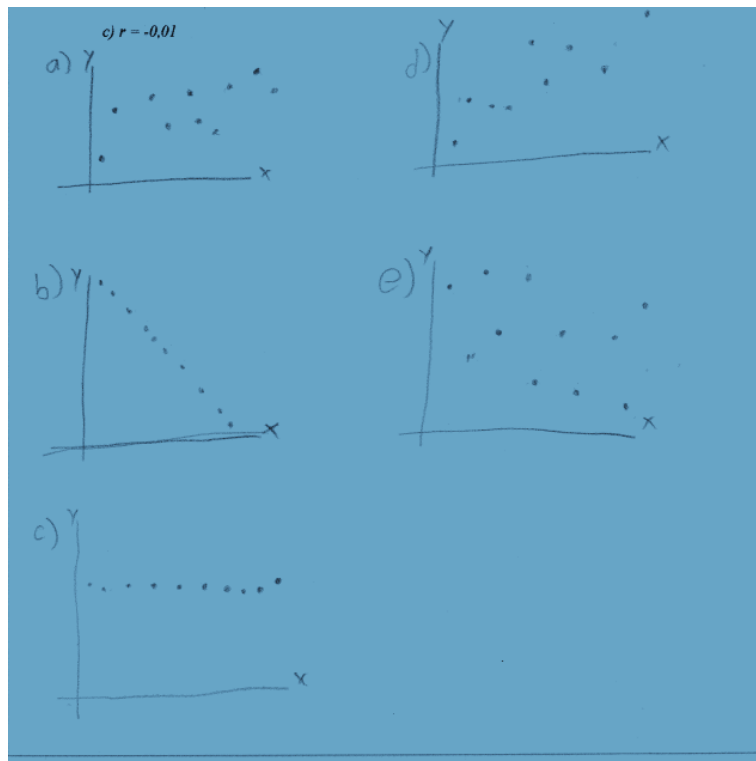


Figura 11.- Actividad 5 elaborada por estudiante 2.

CAPITULO 5

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

La propuesta de este trabajo y su análisis en el capítulo 3, son una aproximación al conocimiento sobre la enseñanza y el aprendizaje de la regresión y la correlación, sirvieron como punto de partida para el análisis en el capítulo 4 de como funcionaron las actividades propuesta en la práctica en el aula y sus consecuencias didácticas, se dio énfasis en el análisis del desarrollo de conocimientos, habilidades y actitudes en los estudiantes relacionadas con la regresión lineal y correlación, implementar las actividades en un ambiente de resolución de problemas cerrando las sesiones con una discusión en grupo da la oportunidad al docente de valorar el desarrollo del conocimiento de los estudiantes e identificar los conceptos que falta reforzar.

5.1 La propuesta didáctica

Para elaborar la propuesta didáctica se basó en el modelo de “Análisis didáctico” propuesto por Pedro Gómez (2002), y se diseñó para estudiantes recién egresados del nivel bachillerato, Se procuró ordenar las actividades para que los estudiantes puedan aumentar sus conocimientos al trabajar cada una de ellas generando ciclos de comprensión, la complejidad de las actividades aumentaba en forma progresiva y se requería en cada una de ellas de la integración de conocimientos previamente abordados, las actividades iniciaron con tratamiento intuitivo donde se revisaron conceptos de estadística descriptiva y representación gráfica de datos con dos variables relacionadas, después a través de la representación gráfica y su análisis hubo una aproximación a la correlación lineal, para finalizar se incorporan los cálculos para hallar el coeficiente de correlación y a los coeficientes de la recta de regresión.

También se puso especial énfasis en que los estudiantes lleguen a darle significado a los conceptos de variable, coeficiente de correlación y darle significado a la recta de regresión. La computadora se utilizó como una herramienta muy importante para el trabajo de la enseñanza procurando que los estudiantes se concentren más en los significados de los resultados que en los procesos de cálculo.

5.2 Aprendizaje de los estudiantes

Al iniciar las actividades de la propuesta didáctica los estudiantes no tenían idea del concepto de regresión lineal y correlación, aunque demostraron saber el significado de promedio, argumentando que en el nivel bachillerato no llevaron la materia de estadística y lo que conocían fue del nivel secundaria. El desarrollo del conocimiento en los conceptos que circularon en el aula fue gradual, en las primeras dos actividades que involucró la resolución de problemas se observa la evolución en la identificación de

variables y la descripción de los datos presentados, también se puede observar en las respuestas a los enunciados que se van introduciendo al concepto de correlación, fue de gran ayuda graficar en Excel y el programa estadístico “R” tomando en cuenta que de manera dinámica pasan de una representación tabular a una representación gráfica lo que hace que el estudiante pueda visualizar y experimentar con el gráfico de dispersión de manera intuitiva la relación entre las variables en juego. En la actividad tres los estudiantes continuaron desarrollando de manera intuitiva la interpretación de los gráficos de dispersión y se fue introduciendo cuando existe relación entre dos variables y como evaluar el grado de relación (correlación lineal). Las siguientes dos actividades exigieron a los estudiantes un mayor esfuerzo en el nivel de comprensión ya que para resolverlas necesitaron de los conocimientos previos mencionados en las actividades anteriores.

Se puede considerar que los estudiantes pudieron alcanzar un nivel de comprensión básico (nivel 1 según el modelo Van Hiele) con respecto al concepto de “regresión lineal”

5.3 Recomendaciones

Haciendo una reflexión acerca de la experiencia obtenida con la aplicación en el aula de la secuencia didáctica “Actividades de aprendizaje para regresión lineal”, se aprecia un aceptable desarrollo del conocimiento del tema en los estudiantes derivado de tres factores principales, el orden de las actividades se seleccionó de tal forma que el desarrollo del conocimiento sea gradual aumentando la complejidad en cada actividad, al final de cada actividad se hacía una discusión en grupo donde los estudiantes expresaban de forma verbal sus reflexiones y el docente podía apreciar el avance, errores y corregir en la siguiente actividad, por último el uso de la computadora para pasar de una representación tabular a una representación gráfica permitió a los estudiantes estar más atentos y visualizar el gráfico de dispersión de forma ordenada.

La recomendación para que el docente que se interese en aplicar esta secuencia obtenga mejores resultados es tener en cuenta y no perder de vista lo siguiente:

- El conocimiento previo del estudiante antes de cada actividad.
- Aumentar la complejidad de las actividades hasta que los estudiantes demuestren que han comprendido el tema en juego.
- La discusión en grupo al cierre de cada sesión es muy importante para la toma de decisiones del docente.
- Es deseable que los estudiantes manejen el programa estadístico que se utilizará durante el curso, se recomienda “R”, pero de no ser así introducirlo en las actividades.
- Por último es importante que el docente maneje al menos lo básico de algún programa estadístico en computadora para utilizarlo como apoyo en las actividades.

REFERENCIAS BIBLIOGRÁFICAS

- Batanero, C., (2001). Didáctica de la Estadística. Departamento de Didáctica de las Matemáticas. Universidad de Granada, España.
- Cabriá, S. (1994). Filosofía de la estadística. Servicio de Publicaciones de la Universidad de Valencia.
- Fernández, F., Monroy, O. y Rodríguez, L. (1998). *Diseño, desarrollo y evaluación de situaciones problemáticas de estadística*. Bogotá, Colombia: Una empresa docente.
- Godino, J. D., Batanero, C. y Cañizares, M. J. (1987). Azar y probabilidad, Fundamentos didácticos y propuestas curriculares, Madrid-
- Gómez, P.(2002).Análisis Didáctico y diseño curricular en matemáticas. Revista EMA, 7(2), 251-292
- Gómez, P. & Lupiáñez, J. L. (2007). Trayectoria hipotética de aprendizaje en la formación inicial de profesores de matemáticas de secundaria. PNA, 1(2) , 79-98
- Jaime, A. y Gutiérrez, A., (1990). “Una Propuesta de Fundamentación para la Enseñanza de la Geometría: El Modelo de Van Hiele” en Teoría y práctica en educación matemática. S. Llinares, M.V. Sánchez Eds. Sevilla: Alfar, 295-384.
- Kahneman, D., Shovic, P. y Tversky, A. (1982). Judgment under uncertainty, Heuristic and biases. New York: Cambridge University Press.
- Lavalle, A., Micheli, E. y Rubio, N. (2006). *Análisis didáctico de regresión y correlación para la enseñanza media*. Relime 9 (3), 383-406
- National Council of Teachers of Mathematics. (2003). Principios para la Educación Matemática. (M. Fernandez, Trad.). España: Sociedad Andaluza de Educación Matemática Thales. (Trabajo original publicado en 2000)
- Rico, L. (1997). *La educación matemática en la enseñanza secundaria*. Barcelona España: ICE-Horsori.
- Simon, M. & Tzur, M. (2004). *Explicating the role of mathematical task in conceptual learning: An elaboration of hypothetical learning trajectory*. Mathematical thinking and learning, 6(2). 91-104.
- Szretter, M. (2013). Apunte de Regresión Lineal. Argentina: Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales.
- Van Hiele, P. M. (1957): el problema de la comprensión (en conexión con la comprensión de los escolares en el aprendizaje de la geometría. Tesis doctoral. Universidad de Utrecht. Utrecht, Holanda.

Duval, R. (2006). Un tema crucial en la educación matemática: La habilidad para cambiar el registro de representación. *La gaceta de la RSME* 9(1), 143- 168.

ANEXO A

SESIÓN 1

ACTIVIDAD 1

Jensen Tire & Auto está por decidir si firma un contrato de mantenimiento para su nueva máquina de alineamiento y balanceo de neumáticos. Los gerentes piensan que los gastos de mantenimiento deben estar relacionados con el uso y recolectan los datos siguientes sobre el promedio de uso semanal (horas) por cada año y gastos anuales de mantenimiento (en cientos de dólares).

Tabla 3.- datos sobre el promedio de uso semanal (horas) por cada año y gastos anuales de mantenimiento (en cientos de dólares).

<i>Uso semanal</i>	<i>Gasto Anual de mantenimiento</i>
13	17
10	22
20	30
28	37
32	47
17	30.5
24	32.5
31	39
40	51.5
38	40

- a) *¿Cuáles son las variables de interés para los gerentes?*
- b) *¿Qué unidad de medida se utiliza para calcular cada variable?*
- c) *¿Qué significa el par ordenado (28; 37)?*
- d) *¿Cuál es el tiempo promedio de uso de la máquina?*
- e) *¿Cuál es el gasto anual medio de mantenimiento?*
- f) *Grafica los puntos de la tabla e indica en el mismo los valores hallados en los incisos c), d) y e).*

SESIÓN 2

ACTIVIDAD 2

Las donaciones de los ex alumnos son una importante fuente de ingresos para las universidades. Si los administradores pudieran determinar los factores que influyen sobre el aumento de la cantidad de alumnos que hacen donaciones, podrían poner en marcha políticas que llevarían a ganancias mayores. Las investigaciones indican que los estudiantes más satisfechos con la relación con sus profesores tienen más probabilidad de titularse, lo que a su vez lleva al aumento de la cantidad de alumnos que hagan donaciones. En la tabla siguiente se muestran datos de 12 universidades de Estados Unidos (American's Best Collage, edición del año 2000). La columna titulada "tasa de titulados" da el porcentaje de alumnos titulados, de los inicialmente inscritos. La columna que tiene como título "% de grupos con menos de 20" da muestra del porcentaje de grupos con menos de 20 alumnos. La columna que tiene como título "Tasa de estudiantes/facultad" da la cantidad total de estudiantes inscritos, dividida entre la cantidad de facultades. Por último, la columna que tiene como título "Tasa de alumnos que donan" da el porcentaje de alumnos que han hecho alguna donación a la universidad.

Tabla 5.- En la tabla siguiente se muestran datos de 12 universidades de Estados Unidos (American's Best Collage, edición del año 2000)

<i>Universidad</i>	<i>Tasa de titulados</i>	<i>% de grupos con menos de 20</i>	<i>Tasa de estudiantes/facultad</i>	<i>Tasa de alumnos que donan</i>
<i>Boston College</i>	85	39	13	25
<i>Brandeis University</i>	79	68	8	33
<i>Brown University</i>	93	60	8	40
<i>California Institute of Technology</i>	85	65	3	46
<i>Carnegie Mellon University</i>	75	67	10	28
<i>Case Western Reserve University</i>	72	52	8	31
<i>College of William and Mary</i>	89	45	12	27
<i>Columbia University</i>	90	69	7	31
<i>Cornell University</i>	91	72	13	35
<i>Dartmouth College</i>	94	61	10	53
<i>Duke University</i>	92	68	8	45
<i>Emory University</i>	84	65	7	37

a) *Elabora tres gráficos de dispersión donde en el eje de las ordenadas de cada uno de ellos figure la tasa de alumnos que donan y en el eje de las abscisas i) tasa de titulados, ii) % de grupos con menos de 20 y iii) tasa de estudiantes/facultad.*

b) *¿Qué puedes decir, en cada caso, sobre el tipo de relación entre las variables consideradas?*

c) *¿Cuál es el promedio de la tasa de titulados? ¿Cuál el de % de grupos con menos de 20? ¿Cuál el de tasa de estudiantes/facultad? Indica en los gráficos.*

d) *¿Cuál es la tasa promedio de los alumnos que donan en estas universidades?*

e) *Indica cuál de los siguientes valores te parece más próximo al coeficiente de correlación lineal del inciso i): -0,98; 0,54; 0,98; 0,73.*

f) *Determina un valor aproximado de r para los incisos ii) y iii).*

SESIÓN 3

ACTIVIDAD 3

Observa los siguientes gráficos y responde:

a) ¿En cuáles de ellos puedes decir que existe relación entre las variables?

b) ¿En cuáles existe relación lineal?

c) En los gráficos que muestran relación lineal entre las variables, analiza el tipo y grado de esa relación y, en cada caso, determina un valor aproximado del coeficiente de correlación.

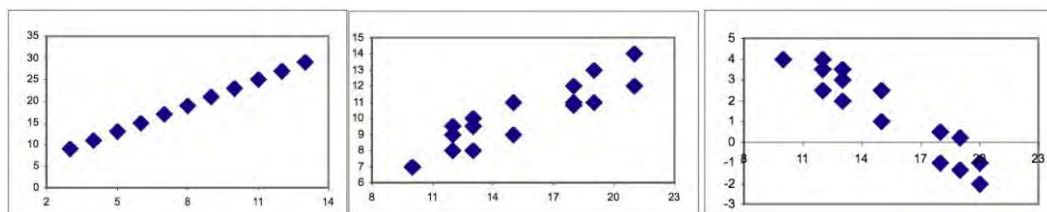


Fig., a

Fig., b

Fig., c

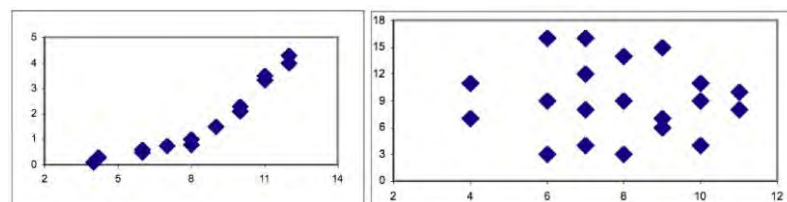


Fig. d

Fig. e

Figura 3.- gráficos de dispersión para la actividad tres

ACTIVIDAD 4

A continuación, se presentan tres afirmaciones referidas a las conclusiones de un estudio acerca de las tasas de nacimiento, suicidio, crecimiento económico y productividad, junto con tres gráficos de dispersión.

Afirmación 1: En países con un desarrollo tecnológico alto, como Japón, Estados Unidos, Alemania, Inglaterra, Francia, Italia y Canadá se tienen bajas tasas de nacimiento (TN) asociadas con altas tasas de suicidio (TS).

Afirmación 2: Algunos economistas afirman que, independientemente de los países que se estudian, a altas tasas de crecimiento (TC) se asocian altas tasas de productividad (TP).

Afirmación 3: Tanto economistas como demógrafos afirman que las tasas de suicidio (TS) no parecen estar correlacionadas con las tasas de productividad (TP). Desafortunadamente, en los gráficos no se colocaron los rótulos de referencia de los ejes. Asocia a cada gráfica una afirmación, completa con el nombre de los ejes y proporciona en cada caso un valor del coeficiente de correlación.



Figura 4.- gráficos de dispersión para la actividad cuatro

ACTIVIDAD 5

Dados los siguientes valores del coeficiente de correlación entre dos variables x , y , dibuja un diagrama de dispersión, con 10 puntos, que se adapte razonablemente a ellos.

- a) $r = 0,25$
- b) $r = -1$
- c) $r = -0,01$
- d) $r = 0,7$
- e) $r = -0,4$

SESION 4

ACTIVIDAD 6

Tomando el enunciado de la actividad uno realiza lo siguiente

- a) *Elaborar un gráfico de dispersión para los datos representados en la tabla del enunciado.*
- b) *Observa el gráfico. Decide si es razonable suponer que existe relación lineal entre las variables; si es así, obtenga la ecuación de regresión estimada que relaciona gastos anuales de mantenimiento con el uso semanal e interpreta los coeficientes. Grafica la recta y estima una aproximación del coeficiente de correlación.*
- c) *Calcula e interpreta el coeficiente de correlación. ¿Estuvo acertada la aproximación que hiciste en el inciso anterior?*
- d) *Jensen piensa que usará la nueva máquina 30 horas a la semana. En ese caso, ¿Qué monto sería el gasto anual de mantenimiento de la empresa?*
- e) *Si el precio del contrato de mantenimiento es \$3000 anuales, ¿recomendaría firmar el contrato de mantenimiento?, ¿Por qué sí o por qué no?*
- f) *Si quisieras estimar el gasto anual de mantenimiento para un uso semanal de la máquina de 5 horas, ¿puedes usar la recta que calculaste en el inciso b? ¿Por qué?*
- g) *¿Qué error se comete en la estimación realizada en el inciso? Indícalo en el gráfico.*

SESION 5

ACTIVIDAD 7

Con esta actividad se propone evaluar el nivel de conocimiento alcanzado por los estudiantes en las sesiones anteriores.

A7.1 Enunciado

Decide si las siguientes sentencias son verdaderas o falsas y justifica brevemente:

- a) Un coeficiente de correlación igual a 1 entre “x, y” significa que “x” está relacionada con “y”, pero un coeficiente de correlación -1 significa que “x” no está relacionada con “y”.*
- b) Si la pendiente de la recta es positiva, el coeficiente de correlación es positivo.*
- c) El signo del coeficiente de correlación depende del signo de la covarianza.*
- d) Un coeficiente de correlación igual a cero indica una no dependencia lineal entre las variables.*
- e) Si existe una correlación alta, se concluye que “x” es causa de “y”.*
- f) Si el coeficiente de correlación es cercano a 1, “x” es directamente proporcional a “y”.*