

## **Análisis didáctico de regresión y correlación para la enseñanza media**

Andrea Lina Lavalle <sup>1</sup>

Elda Beatriz Micheli <sup>1</sup>

Natalia Rubio <sup>1</sup>

### **RESUMEN**

Uno de los importantes campos de aplicación de la regresión lineal y la correlación es el de la enseñanza en la escuela media. Este trabajo se propone hacer un análisis didáctico siguiendo el enfoque propuesto por Fernández, Monroy y Rodríguez en el libro *Diseño, desarrollo y evaluación de situaciones problemáticas de estadística*, donde realizan un estudio desde tres perspectivas: el contenido, la enseñanza y el aprendizaje. En función del análisis, se muestran diversas situaciones de enseñanza que pueden ser útiles para llevar a cabo en las aulas, de manera conjunta con los alumnos.

- **PALABRAS CLAVE:** Regresión lineal, correlación, análisis didáctico.

### **ABSTRACT**

One of the important fields of application of the linear regression and the correlation is that of the teaching in the middle school. This work is proposed to do a didactic analysis following the approach proposed by Fernández, Monroy and Rodríguez in the book *Design, development and evaluation of problematic situations of statistics*, where they carry out a study from three perspectives: the content, the teaching and the learning. In function of the analysis, diverse situations of teaching are shown, that can be useful to carry out in the classrooms, in a joint way with the students.

- **KEY WORDS:** Linear regression, correlation, didactic analysis.

### **RESUMO**

Um dos campos de aplicação mais importantes da regressão linear e da correlação é a educação básica. Este trabalho se propõe a fazer uma análise didática que segue a o enfoque proposto por Fernandez, por Monroy e por Rodriguez no livro, *Esboço, Desenvolvimento e Avaliação De Situações Problemáticas De Estatística*, onde fazem um estudo das três perspectivas: o conteúdo, o ensino e a aprendizagem. Em função

da na análise, apresentamos as diversas situações de ensino que podem ser úteis de se realizar nas salas de aula, da maneira comum com os estudantes.

- **PALAVRAS CHAVE:** Regressão linear, correlação, análise didática.

## RÉSUMÉ

Un des importants domaines d'application de la régression linéaire et de la corrélation est celui de l'enseignement au collège. Ce travail a pour objectif de faire une analyse didactique en suivant l'approche proposée par Fernández, Monroy et Rodríguez dans le livre *Diseño, desarrollo y evaluación de situaciones problemáticas de estadística* (*Conception, développement et évaluation de situations problématiques de statistiques*), où ils réalisent une étude à partir de trois perspectives : le contenu, l'enseignement et l'apprentissage. En fonction de l'analyse, plusieurs situations d'enseignement sont présentées, qui peuvent être utiles en les menant à bien dans les salles de classe de manière conjointe avec les élèves.

- **MOTS CLÉS:** Régression linéaire, corrélation, analyse didactique.

## 1.INTRODUCCIÓN

La educación estadística es un área naciente de la educación matemática que tiene como elementos centrales a la teoría del constructivismo y la resolución de problemas. Algunos principios que están influenciando las prácticas de enseñanza (Burril, 1990) son:

- *El aprendizaje debe ser interactivo y constructivo y se deben generar oportunidades para una discusión creativa, donde cada estudiante pueda poner de su propia parte.*
- *Debe tener lugar la presentación y discusión de puntos de vista conflictivos.*
- *Se debe trabajar hacia un consenso en el cual las ideas estadísticas que hayan sido manejadas sean reconocidas.*
- *Para enseñar los temas tradicionales de la estadística, los estudiantes deben previamente experimentar y trabajar con técnicas sencillas de conteo, tabulación de datos y construcción de gráficas, conjeturar hipótesis y luego verificarlas con métodos estadísticos.*
- *Los temas deben ser presentados bajo formas o diseños que motiven a los estudiantes a ganar experiencia trabajando con datos.*
- *Los proyectos de investigación desarrollados por estudiantes con un fuerte énfasis en la indagación estadística deben ser parte integral de la enseñanza.*
- *El énfasis en cualquier trabajo de estadística debe estar en el análisis y*

*en la comunicación de resultados, no en simples respuestas.*

La regresión y la correlación lineal son contenidos de estadística que ya forman parte del currículum de enseñanza obligatoria. Su inclusión en dicho nivel posibilita el tratamiento estadístico de datos bivariados 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 da hoy día a esta situación es que los valores extremos de una distribución se deben en gran parte al azar.

El análisis que se realiza en este trabajo, de índole preliminar, tiene como fin iniciar una aproximación al conocimiento de la enseñanza y el aprendizaje de la regresión y la correlación y, a través de su estudio, proponer una serie de situaciones de enseñanza que puedan ser relevantes para facilitar la comprensión de tales conceptos.

## 2.FUNDAMENTOS TEÓRICOS Y METODOLÓGICOS

Este trabajo se propone realizar un análisis didáctico, siguiendo el enfoque de Fernández, Monroy y Rodríguez (1998), quienes hacen un estudio desde tres perspectivas: el contenido, la enseñanza y el aprendizaje.

### 2.1 Análisis de contenido

El análisis de contenido abarca la naturaleza del conocimiento estocástico, sus modos de representación y cómo se relacionan con la resolución de problemas. 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. Los *conceptos* describen una regularidad o relación de un grupo de hechos, suelen admitir un modelo o representación y se designan con signos o símbolos. Rico (1997) distingue tres niveles de conocimientos conceptuales: los hechos, los conceptos propiamente dichos y las estructuras conceptuales.

Ahora bien, el *conocimiento procedimental* consiste en los modos de ejecución ordenados de una tarea, de ahí que lo conformen las reglas, algoritmos o procedimientos, de naturaleza claramente secuencial, que se emplean en los procesos de resolución. Por su parte, 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 en la matemática y en particular en la estadística, sus principales manifestaciones son las situaciones –descripciones verbales–, tablas, gráficas y fórmulas. En la siguiente tabla aparecen las formas de representación y su proceso de traducción, que se define como el proceso psicológico involucrado al cambiar de un modo de representación a otro (Fernández et al, 1998).

De <sup>a</sup>	Descripciones	Tablas	Gráficas	Fórmulas
<b>Descripciones</b>		Medir	Dibujar	Modelar
<b>Tablas</b>	Leer		Marcar puntos	Ajustar
<b>Gráficas</b>	Interpretar	Terminación de una lectura		Ajustar curvas
<b>Fórmulas</b>	Reconocer parámetros	Calcular	Dibujar	

A pesar de las limitaciones de esta tabla, resulta útil para dar una idea general sobre el tipo de problemas que se pueden plantear y de las preguntas que pueden surgir en su resolución.

## 2.2 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.

La indagación sobre el libro de texto permite establecer diferencias entre los objetivos de un programa y las formas en que se lleva a la práctica. 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 y calculadoras como herramienta de trabajo.

## 2.3 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.

La identificación de los puntos que presentan dificultad y de los errores detectados en el proceso de enseñanza-aprendizaje (incluso después del mismo) son preocupaciones básicas de la educación estadística. 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.

Según Batanero (2001), la comprensión personal de un concepto es la captación de su significado. Dado que el significado de un objeto se concibe en relación con los contextos institucionales, la comprensión de un concepto por un sujeto, en un momento y circunstancias dadas, implicará la adquisición de los distintos elementos que componen su significado institucional. Por ello, se considera que el análisis de enseñanza, como caracterización del significado institucional del concepto, se encuentra íntimamente relacionado con el aprendizaje. De esta manera, el alumno conoce o comprende un determinado concepto si hay un ajuste entre el significado institucional y el personal que elabora.

### 3. ANÁLISIS DE CONTENIDO EN REGRESIÓN Y CORRELACIÓN

Enseguida aparecen los conceptos que necesitan ser incorporados al tratamiento de la regresión y la correlación lineal para favorecer la comprensión. Para cada concepto se presentan los procedimientos con los que se vincula, así como una serie de relaciones entre conceptos y procedimientos que deben promoverse en el desarrollo de las actividades. También se detallan las formas de representación usuales para estos contenidos.

#### Variable

Es la característica que se estudia u observa en los individuos (o elementos) que conforman el conjunto de unidades a estudiar. En el caso que nos ocupa,

la *regresión lineal simple*, se considera una variable explicativa ( $X$ ) y una variable de respuesta ( $Y$ ); por tanto, las observaciones resultantes son bidimensionales.

#### Datos

Son los valores numéricos asumidos por la variable en cada uno de los elementos que se estudian para que puedan ser comparados, analizados e interpretados. Aquí utilizamos datos bivariados, esto es, pares ordenados que pueden representarse como puntos en el plano cartesiano.

#### Gráfico de dispersión

El gráfico de dispersión o gráfico de la nube de puntos es una representación en un sistema de coordenadas cartesianas de los datos numéricos observados sobre la relación entre dos variables.

#### Recta de regresión

El análisis de regresión consiste en una técnica estadística utilizada para investigar y modelar la relación que hay entre una variable explicativa ( $X$ ) y una variable de respuesta ( $Y$ ). Para poder hacer tal investigación, se debe postular un modelo que está formado por una componente determinística y una componente aleatoria. La primera es una función entre las variables; la segunda considera la función lineal  $\hat{Y}_i = \beta_0 + \beta_1 X_i$ , que representa a la ecuación de la recta en el plano. El coeficiente  $\beta_0$  es la ordenada al origen e indica cuánto vale la variable  $Y$  cuando  $X$  da igual a cero. El coeficiente  $\beta_1$  es la pendiente e indica cuánto varía  $Y$  por cada unidad que aumenta  $X$ . La interpretación de estos coeficientes se efectúa en función de las variables en estudio, es decir, del contexto del problema. Para obtener la recta de

regresión se deben estimar los coeficientes a partir de una muestra de observaciones sobre las variables  $X$ ,  $Y$ . El método comúnmente utilizado es el de *mínimos cuadrados*, donde  $\hat{\beta}_0$  y  $\hat{\beta}_1$  se eligen para minimizar la suma de los cuadrados de los desvíos de las observaciones en torno a la recta 
$$\sum_{i=1}^n [y_i - (\beta_1 x_i + \beta_0)]^2$$

Por ello, se dice que la recta de regresión es la que mejor ajusta a la nube de puntos en el sentido de los mínimos cuadrados.

**Estimación de valores de  $Y$**

Una vez hallada la recta de regresión, es posible reemplazar distintos valores de  $X$  y obtener los valores estimados de  $Y$ . Éstos se definen con la expresión  $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ . Si el valor de  $X$  a reemplazar se encuentra en el rango de valores observados de dicha variable, ocurre una interpolación; si está fuera del rango, hay una extrapolación.

**Coefficiente de correlación lineal**

Mide el grado de relación lineal entre dos variables que varían conjuntamente.

CONCEPTOS	PROCEDIMIENTOS
1. Variables	1.1. Identificar las variables en estudio. 1.2. Distinguir entre una variable respuesta y una variable explicativa. 1.3. Identificar las unidades de medida. 1.4. Calcular la media y el desvío.
2. Datos	2.1. Identificar los valores de las variables como pares ordenados. 2.2. Graficar la nube de puntos. 2.3. Identificar si hay relación entre las variables a partir del gráfico (no existe relación, sí existe relación: lineal, no lineal, etc). 2.3. Observando el gráfico de dispersión de una relación lineal, indicar aproximadamente qué tipo y grado de relación lineal existe (tipo: directa, inversa).
3. Recta de regresión	3.1. Verificar a partir del gráfico la conjetura de relación lineal entre las variables. 3.2. Calcular los coeficientes de la recta. 3.3. Graficar la recta. 3.4. Interpretar los coeficientes de la recta. 3.5. Observando la nube de puntos y la recta, indicar aproximadamente el grado de relación lineal existente.
4. Estimación	4.1. Calcular valores estimados de $y$ . 4.2. Indicar gráficamente el error de estimación. 4.3. Interpretar el error como desvío.
5. Variación conjunta	5.1. Calcular la covariancia. 5.2. Calcular el coeficiente de correlación. 5.3. Interpretar valores del coeficiente de correlación.

**Tabla 1.** Conceptos y procedimientos relacionados con regresión y correlación.

RELACIONES	TIPO
1. Para graficar la nube de puntos hay que tener en cuenta las unidades de medida de cada variable y se debe utilizar una escala adecuada.	1 y 2
2. Los valores de $\bar{x}$ e $\bar{y}$ pueden indicarse en el gráfico.	1 y 2
3. Mediante la nube de puntos o diagrama de dispersión se puede identificar si hay relación lineal entre las variables.	2 y 5
4. Si en la nube de puntos se observa un patrón lineal, las variables presentan una relación lineal.	1 y 3
5. El coeficiente de correlación es una medida sobre el grado de intensidad de la relación lineal existente entre las variables.	1 y 5
6. En la regresión lineal las variables están relacionadas mediante la ecuación de la recta de regresión.	1 y 3
7. A medida que los puntos están más alineados, la correlación entre las variables es más fuerte.	2 y 3
8. La recta de regresión estimada es la que se ajusta lo mejor posible a la nube de puntos (en el sentido de los mínimos cuadrados).	2 y 3
9. El valor estimado de $y$ a partir de los valores de $x$ se obtiene con $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ .	3 y 4
10. La pendiente de la recta indica la variación que se produce en $y$ cuando $x$ varía una unidad; esto debe interpretarse de acuerdo con las variables y el rango de valores estudiados.	1 y 3
11. 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.	1 y 3
12. La recta de regresión pasa por el punto $(\bar{x}, \bar{y})$ .	1 y 3
13. Si $r$ es aproximadamente igual a 0 y los datos tienen un patrón lineal, entonces los puntos caen en una recta horizontal.	3 y 5
14. El coeficiente de correlación auxilia a determinar si hay relación lineal entre las variables.	2 y 5
15. Cuando existe una correlación alta, generalmente los errores de estimación son pequeños.	4 y 5

**Tabla 2.** Relaciones entre los conceptos estadísticos enunciados y procedimientos asociados a regresión y correlación lineal.

## Formas de representación

**Verbal:** Descripción de una situación donde intervienen los conceptos estudiados. Por ejemplo, la descripción de datos bivariados.

**Gráfica:** Empleo de dibujos o diagramas para representar conceptos o procedimientos. En este caso, se utiliza el diagrama de dispersión y la gráfica de la recta de regresión estimada.

**Simbólica:** Manejo de la simbología para expresar nociones o conceptos estadísticos. Por ejemplo,  $\bar{y}$ ,  $\bar{x}$ ,  $\hat{y}$ .

**Tabular:** Se utiliza para representar el conjunto de pares de valores numéricos de datos bivariados.

## 4. ANÁLISIS DE ENSEÑANZA EN REGRESIÓN Y CORRELACIÓN

Se analizaron los libros de texto de matemáticas donde se abordan los contenidos de estadística, particularmente los tocantes a la regresión y correlación. Dichos libros los propone el Ministerio de Educación en Argentina para su uso en la enseñanza de las matemáticas en el nivel medio y se encuentran disponibles en las bibliotecas de las escuelas. Al estudiarlos, nos centramos en el enfoque con que se presentan los temas de regresión y correlación, el nivel de profundidad con que son abordados, si es que se deducen las fórmulas de la regresión y correlación, así como si se sugiere el uso de herramienta tecnológica.

Los libros de matemáticas analizados fueron:

### 1. *Matemática 9:*

Camus, N. (2000). Libro de texto propuesto

para ser utilizado con alumnos de 13-14 años.

### 2. *Matemática 2 Polimodal:*

Fones, M. A. (1998). Libro de texto propuesto para alumnos de 16 años.

### 3. *Matemática 4:*

Turner, S. (2000). Libro de texto propuesto para ser utilizado con alumnos de 16 años.

### 4. *Matemática 4:*

Barallobres, S. (1998). Libro de texto propuesto para ser utilizado con alumnos de 16 años.

### 5. *Matemática 1 Bachillerato:*

De Guzmán, S. (1999). Libro de texto propuesto para ser utilizado con alumnos de 13 a 15 años.

### 6. *Matemática 3 Bachillerato:*

De Guzmán, S. (1999). Libro de texto propuesto para ser utilizado con alumnos de 15 a 16 años.

### 7. *Matemática II COU:*

De Guzmán, S. (1999). Libro de texto propuesto para la preparación al ingreso a la universidad.

## 4.1 Enfoque de la presentación del tema

En todos los textos, excepto en el 3, se introduce el tema a través de problemas. Muestran una intención explícita por desarrollar los conceptos, haciendo énfasis en la construcción crítica de las habilidades del pensamiento, no en el enfoque tradicional de la enseñanza de la estadística, donde se pone de manifiesto



el interés por el manejo de fórmulas y ecuaciones.

El texto 1 contiene un problema motivador, en el que trata de llegar a los conceptos sólo de forma intuitiva, sin definiciones ni fórmulas. Aquí no hay una intención visible de institucionalización, es una introducción en el tema que incluye los conceptos de correlación y de recta de regresión. Los textos 2, 4, 5, 6 y 7 también presentan el tema mediante tablas de datos, que indican diferentes tipos de relaciones entre variables y luego plantean el análisis de las distintas nubes de puntos asociadas a dichas tablas.

En el texto 3 hay ejemplos resueltos y a continuación se proporcionan las definiciones con sus respectivas fórmulas. Luego, se proponen ejercicios muy similares a los de los ejemplos. Tal enfoque es del tipo ejemplo-teoría-ejercitación.

## 4.2 Nivel de profundidad y temas abarcados

En el texto 1 se trabaja de forma intuitiva, sin llegar a formalizar ningún concepto ni realizar cálculos. La aproximación al concepto de relación entre variables se hace a través de gráficos y tablas de datos.

El texto 2 lleva a cabo un enfoque a profundidad de los temas abarcados, realizando aspectos significativos para la comprensión, mientras que el 3 tiene el mismo alcance que el texto 2 en cuanto a los contenidos, mas solamente utiliza ejemplos y fórmulas.

En el texto 4 se trabaja el tema de correlación y una aproximación intuitiva al concepto de regresión lineal. El 5, 6 y 7 desarrollan los conceptos de correlación, recta de regresión lineal y coeficiente de correlación, con una muy buena selección de problemas y un enfoque orientado a la comprensión.

Conceptos y procedimientos	1	2	3	4	5	6	7
Gráfico de dispersión o nube de puntos	X	X	X	X	X	X	X
Valor aproximado de $r$	X	X			X	X	X
Covariancia				X		X	X
Cálculo de $r$ . Fórmula		X	X	X		X	X
Interpretación de $r$		X	X	X	X	X	X
Interpretación del diagrama de dispersión	X	X	X	X	X	X	X
Gráfico de la recta de regresión		X	X			X	X
Ecuación de la recta de regresión		X	X			X	X
Centro de gravedad		X		X		X	X
Alusión a las relaciones causa-efecto	X	X					X
Alusión a las regresiones no lineales		X					X
Estimación de $y$	X	X	X		X		X
Condiciones para realizar estimaciones de $y$		X			X		X
Relaciones sin sentido		X					

**Tabla 3.** Conceptos y procedimientos abarcados en los diferentes textos analizados.

4.3 Deducción de fórmulas

Sólo en el texto 6, que está destinado a los últimos años de enseñanza secundaria, se deduce la fórmula de cálculo de la covarianza aplicando propiedades de sumatoria. Además, se explica que el método para obtener la pendiente y la ordenada al origen de la recta de regresión tiene como base minimizar la expresión  $\sum [y - (\beta_1 \cdot x_i + \beta_0)]^2$ . Si bien la deducción de la fórmula de cálculo no se desarrolla, la explicación del método haya sustento en un gráfico.

4.5 Uso de las computadoras como herramientas de trabajo

Ninguno de los textos propone algún tipo de actividad para el uso de computadoras.

4.6 Conclusión del análisis realizado

En general, los libros de texto analizados dan mayor importancia al análisis de correlación lineal, lo que refuerzan con el uso de diferentes diagramas y solicitando a los alumnos que analicen el tipo y grado de relación lineal. Con respecto al

	1	2	3	4	5	6	7
Interpretar tipo de relación lineal (directa,inversa o nula) a partir de la nube de puntos					X	X	
Identificar el grado de relación lineal	X			X	X	X	X
Proponer un valor aproximado de $r$ a partir de gráfico	X	X			X	X	X
Cálculo de $r$		X	X	X		X	X
Interpretar $r$		X		X		X	
Graficar la nube de puntos							X
Gráficar la recta de regresión		X			X	X	X
Calcular la ecuación de la recta de regresión		X	X			X	X
Estimar $y$ para valores de $x$		X	X		X	X	X
Identificar condiciones para realizar una estimación		X			X	X	X
Búsqueda de datos: mediciones		X					
Utilización de datos reales (diarios, revistas, TV, anuarios estadísticos, internet)		X			X	X	X
Identificación de posibles relaciones estadísticas							X

Tabla 4. Tipo de situaciones problemáticas propuestas en los textos analizados.

4.4 Tipos de situaciones problemáticas

En las situaciones problemáticas que proponen los textos analizados hay una mayor cantidad de actividades donde se utiliza más la relación lineal directa que la inversa. Como detalla Estepa (1995), algunos estudiantes tienen dificultades con la relación lineal inversa, por lo cual es recomendable trabajar con la misma intensidad ambas relaciones.

tratamiento destinado a los alumnos de 16 años (cuarto año de secundaria), abarca más temas y mayor nivel de profundidad, mientras que para los alumnos de menor edad se propone una aproximación intuitiva.

Los textos que piden el cálculo de la recta de regresión emplean la fórmula punto

pendiente, la cual es conocida por los alumnos desde segundo o tercer año de la escuela secundaria. Así, el contenido puede relacionarse con los conocimientos previos de los estudiantes.

## 5. ANÁLISIS DE APRENDIZAJE EN REGRESIÓN Y CORRELACIÓN

El estudio acerca del razonamiento sobre la asociación estadística en general fue iniciado por Piaget e Inhelder (1955), quienes utilizaron solamente variables dicotómicas (tabla de contingencia de 2x2) y analizaron la habilidad de los alumnos a partir de los 12 ó 13 años para diferenciar correlación directa, inversa e independencia, así como las estrategias que emplean para establecer sus juicios. Los autores consideraron que la comprensión de la idea de asociación implicaba las de proporción y probabilidad. Por ello, sugirieron la enseñanza de los conceptos relacionados con la asociación estadística en la etapa de las operaciones formales.

### 5.1 Conocimientos previos

El tratamiento que se propone realizar a los temas de regresión y correlación en el nivel de enseñanza secundaria es de tipo descriptivo. No se pretende la comprensión de los supuestos del modelo ni de la inferencia estadística, de ahí que los conocimientos previos se refieran a los siguientes contenidos:

- **Conceptos matemáticos:** Puntos en el plano, variable, función, función lineal, ecuación de la recta, pendiente, ordenada al origen, representación gráfica de la recta, ecuación de la recta que pasa por un punto con pendiente conocida.
- **Conceptos estadísticos:** Datos

bivariados, media, desvío estándar y variabilidad.

### 5.2 Errores y dificultades de aprendizaje

En general, las investigaciones tanto de tipo didáctico como psicológico acerca de la regresión y la correlación son escasas, a pesar de su importancia en la estadística y en el currículum de matemática.

Hay errores que atañen al cálculo de los coeficientes de la recta y del coeficiente de correlación; otros, a las representaciones gráficas (Fernández & Monroy, 1995). Con frecuencia se observa la omisión o mal manejo de las escalas en uno o ambos ejes y la ausencia del sentido de agrupación o de escala.

Por otra parte, la destreza en la lectura crítica de datos es una componente de la alfabetización cuantitativa y una necesidad en nuestra sociedad tecnológica. Batanero et al. (1994) describen los tres niveles distintos de comprensión de los gráficos establecidos anteriormente por Curcio:

**a) Leer los datos:** Este nivel de comprensión requiere de una lectura literal del gráfico; no se interpreta la información que contiene. En una nube de puntos, dicho nivel se refiere a cuestiones sobre la lectura de las escalas o encontrar el valor de una de las coordenadas de uno de los puntos, dado el valor de la otra coordenada.

**b) Leer dentro de los datos:** Incluye la interpretación e integración de los datos en el gráfico; además, requiere de la habilidad para comparar cantidades y el uso de otros conceptos y destrezas matemáticas. Le conciernen aspectos sobre la intensidad de la covariación, si la relación podría ser representada

mediante una función lineal o si la dependencia es directa o inversa.

**c) Leer más allá de los datos.** Requiere que el lector haga predicciones e inferencias a partir de los datos sobre informaciones que no se reflejan directamente en el gráfico. Por ejemplo, la predicción del valor de la coordenada  $y$  para un valor de la coordenada  $x$ .

Respecto a la tarea de determinar un **valor aproximado del coeficiente de correlación**, distintos investigadores (Sánchez Cobo et al., 2000) han reportado que mejora cuando hay más datos y la correlación es alta.

Por otro lado, la estimación resulta más exacta cuando los datos se dan en la forma gráfica que en la tabular. En general, cuando hay teorías previas acerca de las variables consideradas se sobreestima la correlación, mientras que cuando no existen es necesaria una correlación fuerte para que los estudiantes la detecten (Estepa & Sánchez Cobo, 2001).

Batanero et al. (1998) mencionan que otro tipo de dificultades relacionadas con la **interpretación** de la correlación son:

- **La concepción determinista:** Cuando el estudiante sólo admite para un valor de la variable independiente un único valor de la variable dependiente.
- **La concepción local:** Cuando el estudiante se basa en parte de los datos presentados y no en todos. Si la parte muestra un tipo de asociación, adopta ese tipo para todo el conjunto de datos.

En el caso particular de la regresión y la correlación lineal, se ha encontrado influencia de las concepciones

determinista, local y causal en los juicios de asociación (Estepa & Sánchez Cobo, 2001).

Las **teorías previas** que los alumnos tienen respecto a las variables involucradas afectan a menudo el juicio de asociación –concepción causal– y la idea intuitiva del grado de relación entre las variables. Sánchez Cobo et al. (2000), en su estudio sobre la manera en que las personas proporcionan un valor aproximado de la correlación a partir de diferentes representaciones, reportan que no existe diferencia en los errores detectados en esta tarea, según haya teorías previas a favor o en contra. Sin embargo, la influencia de teorías previas se reporta en los casos de concordancia y dependencia indirecta.

Ahora bien, se ha detectado que algunos estudiantes confunden la variación conjunta y la proporcionalidad (relación directa y directamente proporcional); en la interpretación del valor del coeficiente de correlación no distinguen entre la intensidad y el sentido de la correlación –por ejemplo, cuando se dice: “existe una relación directa porque el coeficiente de correlación es alto”–, así como entre asociación y causalidad. Aquí, Estepa (1995) señala que se deben proponer actividades que hagan comprender al alumno que la causalidad implica asociación, pero no a la inversa.

Con respecto a las actividades de **predicción**, se distinguen dos tipos de dificultades. Por un lado, las relacionadas con la interpolación (predicción para valores dentro del rango), como la falta de interpretación de los resultados en relación con contexto del problema; un ejemplo es discretizar el resultado. Por otro lado, las asociadas con la extrapolación (predicción para valores fuera del rango).

## 6. ORIENTACIONES PARA EL DISEÑO DE ACTIVIDADES PARA LA ENSEÑANZA

A través del análisis didáctico sobre los contenidos en regresión y correlación se ofrecen algunas sugerencias descriptivas que orienten la elección de situaciones de enseñanza destinadas al nivel secundario y favorezcan la comprensión de los contenidos.

Por ello, se propone trabajar los contenidos formalmente a partir de los 15 ó 16 años (periodo de operaciones formales). Tal sugerencia se basa en que entre el segundo y el tercer año de escuela secundaria (14 a 15 años) los alumnos incorporan los conceptos de función, función lineal y sistemas de ecuaciones lineales.

En principio, se propone hacer un tratamiento intuitivo que comprenda la revisión de conceptos relacionados con la estadística descriptiva y la representación gráfica de datos bivariados. Luego se sugiere una aproximación a la correlación lineal a través del análisis de gráficos y, por último, se incorporan los cálculos referidos al coeficiente de correlación y a los coeficientes de la recta de regresión, con sus respectivas interpretaciones, así como la utilización de la recta ajustada para la estimación de valores de  $y$ . Asimismo, realizar gráficos con distintas nubes de puntos, trazar para cada una de ellas su recta de regresión, ver similitudes y diferencias, analizar qué ocurre si se cambian las escalas.

El uso de las calculadoras y las computadoras es un componente muy importante que se relaciona con el trabajo en la enseñanza de la estadística. Con la nueva tecnología es posible acceder de

forma rápida a representaciones gráficas que hacen mucho más interesante y efectivo el análisis de datos para los estudiantes, brindándoles la posibilidad de concentrarse más en el significado de los resultados que en los procesos de cálculo. Por ello, se propone trabajar con los utilitarios de fácil acceso y que se encuentran en los equipos informáticos de las escuelas. Se puede mencionar el *MS Excel*, que tiene un módulo interesante de análisis de datos, y el *INFOSTAT* (Universidad Nacional de Córdoba, Argentina). Hay una amplia variedad de paquetes estadísticos con distintos niveles de adecuación al nivel de maduración del alumno en secundaria, los cuales son más difíciles de acceder por su costo.

A continuación, se realiza un análisis didáctico sobre diferentes propuestas de actividades para la enseñanza. La elección está sustentada en los textos analizados y en la experiencia de las autoras en cursos de formación y actualización docente. El enfoque presentado y los incisos que conforman las actividades están presentados en tres grupos, de acuerdo con los contenidos que involucran:

- 1) Tratamiento de datos bivariados
  - Variables
  - Descripción de las variables: unidad de medida, media y desvío estándar de cada una
  - Gráfico de la nube de puntos
  - Ubicación de  $\bar{x}$  e  $\bar{y}$  en el gráfico
- 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, utilizando

$$\hat{\beta}_1 = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \quad \text{y} \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

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

El *análisis de contenido* explicita los conceptos y procedimientos que se ponen en juego en la actividad, las interrelaciones

entre ellos y las representaciones externas que se pueden utilizar; el *análisis de enseñanza* detalla el nivel de profundidad, si se necesita entrenamiento previo en algún tipo de procedimiento, la intención de algunas preguntas propuestas y posibles ampliaciones, mientras que el *análisis de aprendizaje* evalúa los posibles errores que pueden surgir en la puesta en marcha de la actividad, los conocimientos previos para abordarla y el tipo de lectura de datos que se solicita.

●

## 7. TRATAMIENTO DE DATOS BIVARIADOS

### Actividad 1

*El entrenador de un equipo de atletismo está controlando los tiempos registrados por su equipo para la carrera de 100 metros planos. Elaboró la siguiente tabla, que describe los tiempos logrados y las horas diarias de entrenamiento para uno de sus atletas:*

Horas diarias de entrenamiento	1	2	3	4	5	6
Tiempo (en segundos)	12	11,8	11,2	11	10,8	10,2

- a) ¿Cuáles son las variables de interés para el entrenador?
- b) ¿Qué unidad de medida se utiliza para calcular cada variable?
- c) ¿Qué significa el par ordenado (3; 11,2)?
- d) ¿Cuál es el tiempo promedio de entrenamiento?
- e) ¿Cuál es el tiempo medio logrado por este atleta en los 100 metros planos?
- f) Grafica los puntos de la tabla e indica en el mismo los valores hallados en los incisos c), d) y e).

## **Análisis de contenido**

Esta actividad es una modificación de la presentada en el libro de texto analizado como libro 1.

Los conceptos y procedimientos estudiados se refieren a:

Tabla 1. Conceptos y procedimientos: 1.1- 1.3- 1.4- 2.1- 2.2

Las interrelaciones entre los conceptos que se mencionan se refieren a:

Tabla 2. Relaciones: 1- 2

Por otro lado, esta actividad permite pasar de la forma de representación tabular a la gráfica y de la forma tabular a la descripción mediante la interpretación.

## **Análisis de enseñanza**

Las actividades de tratamiento de datos bivariados se proponen como introducción, a fin de permitir que el alumno se familiarice con este tema. Por eso el nivel de profundidad es elemental, ya que propone la vinculación con conceptos tratados con anterioridad, como el gráfico de pares ordenados y el cálculo de promedios. Si no se ha hecho junto con los contenidos de estadística descriptiva, es necesario entrenar a los alumnos en el uso de la calculadora.

Se sugiere que el trabajo sea en forma grupal y que el docente identifique errores a partir de los resultados de los alumnos, con el fin de conocer la necesidad de profundizar en actividades similares.

Como actividad complementaria, se propone la realización de un proyecto de investigación interdisciplinario junto con

el profesor de Educación Física para obtener datos reales que puedan ser analizados de la misma manera.

Por otro lado, esta actividad puede ser retomada para ofrecer un valor aproximado del coeficiente de correlación y luego su cálculo, al igual que para obtener la recta de regresión. De esta manera se puede completar con los demás niveles de comprensión de los gráficos.

Ejercicios de esta índole pueden encontrarse en los libros 1 y 2. En el resto, las actividades introductorias presentadas involucran solamente el cálculo de la media, el desvío estándar y el gráfico de la nube de puntos.

## **Análisis de aprendizaje**

En el desarrollo de esta actividad pueden surgir errores al elegir la escala en el eje vertical, en el cálculo de los promedios y en la interpretación de pares ordenados.

Ahora bien, el hecho de solicitar el cálculo de los promedios de cada variable permite relacionar con los conocimientos previos de estadística descriptiva. La construcción del gráfico hace que el alumno se familiarice con el uso de diferentes escalas.

Tanto en el gráfico como en los incisos b) y c) se realiza una lectura de los datos conforme al nivel a) de comprensión, enunciado por Curcio.



## **8.RELACIÓN ENTRE VARIABLES-CORRELACIÓN**

### **Actividad 2**

*La siguiente tabla contiene las posiciones correspondientes al resultado final del torneo Clausura 2006 del fútbol argentino:*

POS.	EQUIPO	PTS	Jugados	Ganados	Empates	Perdidos	Goles a Favor	Goles en Contra
1	Boca	43	19	13	4	2	37	14
2	Lanús	35	19	10	5	4	26	15
3	River	34	19	9	7	3	39	24
4	Gimnasia (J)	33	19	10	3	6	26	16
5	Gimnasia	32	19	9	5	5	31	22
6	Newell's	31	19	8	7	4	27	18
7	Banfield	31	19	10	1	8	28	22
8	San Lorenzo	28	19	7	7	5	15	14
9	Olimpo	26	19	7	5	7	22	22
10	Vélez	25	19	5	10	4	21	18
11	Estudiantes	24	19	6	6	7	25	31
12	Independiente	23	19	6	5	8	18	18
13	Rosario Central	23	19	5	8	6	15	17
14	Argentinos	22	19	5	7	7	26	28
15	Arsenal	22	19	5	7	7	17	20
16	Colón	20	19	5	5	9	19	26
17	Quilmes	20	19	5	5	9	16	28
18	Racing	19	19	5	4	10	14	24
19	Instituto	13	19	3	4	12	16	37
20	Tiro Federal	12	19	3	3	13	15	39

a) *Elabora tres sistemas ortogonales donde en el eje de las ordenadas de cada uno de ellos figuren los puntos obtenidos por equipo y en el eje de las abscisas la cantidad de: i) partidos ganados, ii) partidos perdidos y iii) empatados, respectivamente.*

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

c) *¿Cuál es el promedio de partidos ganados? ¿Cuál el de perdidos? ¿Cuál el de empatados? Indica en los gráficos.*

d) *¿Cuál es el número promedio de goles a favor y en contra de este torneo?*

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).*

### Análisis de contenido

Los conceptos y procedimientos analizados en esta actividad se refieren a:

Tabla 1. Conceptos y procedimientos: 1.4- 2.2- 2.3- 2.4.

Las interrelaciones entre los conceptos que se mencionan aluden a:

Tabla 2. Relaciones: 2 - 3 - 4 - 5 - 7.

Esta actividad permite pasar de la forma tabular a la gráfica analizando, en cada caso, el tipo de relación entre las variables consideradas. Además, hace que se transite de la forma gráfica a la descripción a través de la interpretación



sobre la relación dada. En las tres gráficas pedidas se representan tres tipos diferentes de relación.

### **Análisis de enseñanza**

En esta actividad se propone que el alumno identifique distintos tipos de relación mediante el gráfico de nubes de puntos. Pueden surgir dificultades en el del inciso a), donde tenga que intervenir el docente, mientras que el propósito de los incisos c) y d) es retomar contenidos de estadística descriptiva. En la asignación de un valor de correlación, se avanza respecto a actividades como la anterior. Para ello, es necesario que los alumnos conozcan el recorrido del coeficiente y su interpretación en relación con su magnitud y signo.

La actividad se puede retomar después para que los alumnos calculen el valor del coeficiente de correlación y lo comparen con los valores sugeridos en los incisos e) y f). Otros ejercicios de este tipo pueden llevarse a cabo como proyectos de investigación a cargo de los alumnos, utilizando diferentes tablas de datos. Por ejemplo, en la página de internet [www.amstat.org/publications/jse/](http://www.amstat.org/publications/jse/) se pueden hallar conjuntos interesantes de datos para analizar. Aunque puede ser más relevante que los alumnos consigan tablas de posiciones sobre los equipos de su interés.

En los libros analizados hay actividades similares que incluyen por lo general el gráfico de la nube de puntos y la valoración de la correlación a partir del mismo. Aquí, el libro más completo es el 6.

### **Análisis de aprendizaje**

En el desarrollo de esta actividad se pretende que los alumnos puedan hacer una lectura de los datos (nivel a); 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 la nube de puntos.

A partir de los incisos e) y f) se espera que los alumnos superen posibles errores en la aproximación del coeficiente de correlación, los cuales estriban en la dificultad de relacionar el signo negativo del coeficiente con relación inversa entre las variables y de proporcionar valores aproximados del coeficiente de correlación lineal en nubes de puntos poco densas.

### **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.*

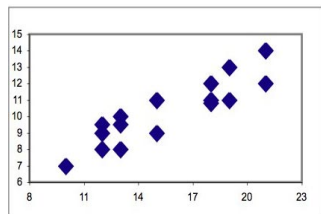


Figura a)

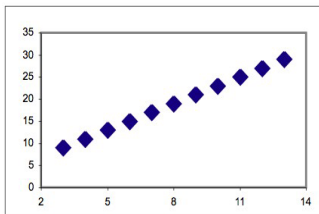


Figura b)

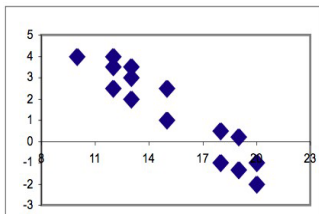


Figura c)

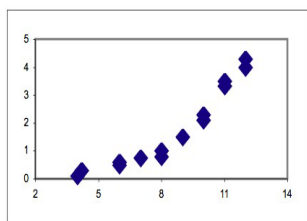


Figura d)

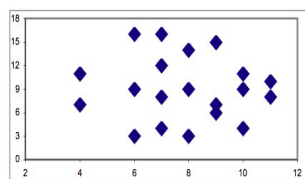


Figura e)

### Análisis de contenido

Los conceptos y procedimientos analizados en esta actividad se refieren a:

Tabla 1. Conceptos y procedimientos: 2.2 – 2.3 – 2.4.

Las interrelaciones entre los conceptos conciernen a:

Tabla 2. Relaciones: 3 – 4 – 5 – 7 – 13.

La actividad permite identificar relaciones entre variables, a través del análisis de gráficos de distintas nubes de puntos, e involucra distintos tipos de relación lineal, una relación no lineal y un gráfico de no relación. Por otro lado, es posible pasar de la representación gráfica a la descripción mediante la interpretación.

### Análisis de enseñanza

Es una actividad intuitiva, ya que no pide cálculos, sino el análisis de representaciones gráficas y la determinación de valores

aproximados de los coeficientes de correlación. Por ello, puede realizarse en grupos y luego proponer una discusión sobre las respuestas con todos los alumnos; aquí, el docente debe guiar la discusión y evaluar las aproximaciones de  $r$ .

Para alentar el uso de computadoras, los datos podrían darse en forma tabular y proponer a los alumnos que utilicen un paquete informático que les posibilite realizar el análisis propuesto. En todos los libros analizados hay actividades similares; sin embargo, en ninguno se plantea el uso de computadoras.

### 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.

### 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.*

### Análisis de contenido

Los conceptos y procedimientos analizados en esta actividad son:

Tabla 1. Conceptos y procedimientos: 1.1 - 1.2 - 2.4.

Las interrelaciones entre los conceptos que se mencionan aluden a:

Tabla 2. Relaciones: 3 - 4 - 5 - 7.

Esta actividad, que es una modificación de la presentada por Fernández et al. (1998), propone trabajar con las formas de representación verbal y gráfica. Con ello, se pretende que el alumno reconozca la relación establecida en forma textual e identifique la figura correspondiente.

### Análisis de enseñanza

Si bien la 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.

Ninguno de los libros analizados propone la relación entre representaciones gráficas y enunciados verbales, pero sí hay actividades donde se pide realizar una aproximación del grado de relación para distintos pares de variables.

Gráfico 1

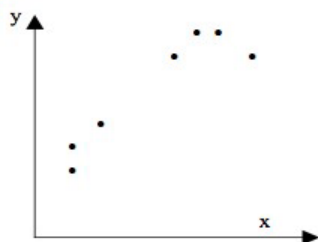


Gráfico 2

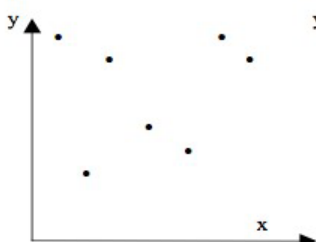
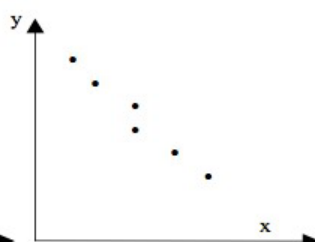


Gráfico 3



## Análisis de aprendizaje

Si bien los conocimientos previos necesarios no son diferentes de las actividades anteriores, es preciso que los alumnos interioricen el significado de indicadores estadísticos como las tasas de nacimiento, suicidio, productividad o crecimiento económico. Es posible que surjan errores asociados con la falta de interpretación de dichos indicadores.

El tipo de lectura e interpretación pedida en este problema requiere de una comprensión tanto de los gráficos que involucra la lectura dentro de los datos (nivel b), como del tipo de relación enunciado verbalmente.

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

### Análisis de contenido

Los conceptos y procedimientos analizados en esta actividad conciernen a:

Tabla 1. Conceptos y procedimientos: 2.2 – 5.3.

Las interrelaciones entre los conceptos se refieren a:

Tabla 2. Relaciones: 3 – 4 – 5 – 7 – 13.

Esta actividad, que fue propuesta por Sánchez Cobo et al. (2000), tiene como propósito que el alumno traduzca la información brindada por una fórmula a un diagrama.

## Análisis de enseñanza

La actividad exige un tratamiento previo acerca del coeficiente de correlación. Por ejemplo, hacer previamente el proceso inverso: dada una nube de puntos, 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.

Esta actividad podría completarse solicitando a los alumnos que, para cada inciso, describan dos situaciones reales donde sea razonable obtener el coeficiente de correlación en función del tipo de dependencia entre las variables y su intensidad (Sánchez Cobo, 2000). Con tal agregado, se pondrían en juego teorías previas.

Actividades de esta índole se presentan en todos los libros analizados, excepto en el 3 y el 4.

### 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.

## 9.REGRESIÓN

### Actividad 6

*El responsable de un camping en una zona turística del país ha observado, en las diez*

*semanas que ha permanecido abierto, la temperatura media de cada semana ( $x$ ) y la cantidad de gaseosas consumidas ( $y$ ). La información que obtuvo es la siguiente:*

$x$	20	38	22	41	40	39	34	15	19	25
$y$	200	350	240	420	400	400	340	150	190	280

*a) Elige una escala adecuada para graficar la información.*

*b) Observa el gráfico. Decide si es razonable suponer que existe relación lineal entre las variables; si es así, calcula la recta de regresión e interpreta los coeficientes. Grafica la recta.*

*c) Calcula e interpreta el coeficiente de correlación. ¿Estuvo acertada la aproximación que hiciste en el inciso a)?*

*d) ¿Cuántas gaseosas se consumirán si la temperatura media es de 15 grados?*

*e) Si quisieras estimar la cantidad de gaseosas consumidas para un día en el que la temperatura media fuera de 5°, ¿puedes usar la recta que calculaste en el inciso b)? ¿Por qué?*

*f) ¿Qué error se comete en la estimación realizada en el inciso? Indícalo en el gráfico.*

### **Análisis de contenido**

Los conceptos y procedimientos analizados en esta actividad son:

Tabla 1. Conceptos y procedimientos: 1.1 – 2.2 – 2.4 – 3.1 – 3.2 – 3.3 – 3.4 – 4.1 – 4.2 – 4.3 – 4.4 – 5.1 – 5.3.

Las interrelaciones entre los conceptos se refieren a:

Tabla 2. Relaciones: 1 – 3 – 6 – 8 – 9 – 10 – 11 – 14.

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. Los incisos b) y c) plantean pasar de la tabla a las fórmulas mediante el ajuste correspondiente.

La actividad comienza a nivel intuitivo, pero permite comprobar las aproximaciones a través del cálculo de los coeficientes correspondientes. En el inciso e) se pretende que los alumnos identifiquen cuándo es posible hacer una estimación que involucre una extrapolación.

### **Análisis de enseñanza**

Esta actividad, que reúne todos los conceptos y procedimientos de las anteriores, puede desarrollarse 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.

En los libros analizados se presentan actividades similares, excepto en el 1 y el 4. Sin embargo, ninguno pide el cálculo del error de estimación. Los textos 2, 5, 6 y 7 contienen actividades sobre análisis de extrapolaciones.

### **Análisis de aprendizaje**

Esta actividad, que relaciona los conceptos anteriores, 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.

### Actividad 7

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

*a) Un coeficiente de correlación igual a 1 entre  $x$  y  $y$  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  $y$ .*

### Análisis de contenido

Esta actividad plantea que se haga un

tratamiento conceptual de los contenidos, permitiendo la discusión de los temas tratados y de puntos de vista conflictivos.

### Análisis de enseñanza

La actividad sirve como autoevaluación de conceptos, por lo cual es recomendable que se realice en forma individual, con una instancia posterior de puesta en común y discusión.

### Análisis de aprendizaje

En la primera afirmación se indaga acerca de la concepción unidireccional de la asociación; en el inciso e) se analiza la concepción causal y en el f) se evalúa si el alumno confunde la variación conjunta con la proporcionalidad.



## 10. CONSIDERACIONES FINALES

El análisis y la propuesta detallada en este trabajo constituyen una aproximación al conocimiento sobre la enseñanza y el aprendizaje de la regresión y la correlación. La complejidad asociada al significado y comprensión de estos contenidos pone de relieve la necesidad de investigar acerca de cuáles son los procesos de enseñanza que favorecen un aprendizaje significativo para los alumnos. Por ello, se considera que este estudio es también un punto de partida para el análisis de cómo funcionan las actividades propuestas en la práctica áulica y sus consecuencias didácticas.

## ●

## BIBLIOGRAFÍA

Barallobres, S. (1998). *Matemática 4*. Buenos Aires, Argentina: Aique.

Batanero, C. (2001). *Didáctica de la estadística*. Granada, España: Grupo de Investigación en Educación Estadística. Obtenido de la página electrónica [www.ugr.es/~batanero/](http://www.ugr.es/~batanero/).

Batanero, C.; Godino, J. y Estepa, A. (1998). Construcción del significado de la asociación estadística mediante actividades de análisis de datos. *Proceedings of the 22<sup>nd</sup> Conference of the International Group for the Psychology of mathematics Education* (Vol. 1, pp. 221-236). South Africa: University of Stellenbosch.

Batanero, C.; Godino, J.; Green, D.; Holmes, P. y Vallecillos, A. (1994). Errores y dificultades en la comprensión de los conceptos estadísticos elementales. *International Journal of Mathematics Education in Science and Technology* 25 (4), 527-547.

Burril, Gail. (1990). Quantitative literacy: leadership training for masters teachers. En Anne Hawkins (Ed.), *Training teachers to teach statistics* (pp. 219-227). The Netherlands: International Statistical Institute-UNESCO.

Camus, N. (2000). *Matemática 9*. Buenos Aires, Argentina: Aique.

De Guzmán, S. (1999). *Matemática 1. Bachillerato*. Madrid, España: Anaya.

De Guzmán, S. (1999). *Matemática 3. Bachillerato*. Madrid: España: Anaya.

De Guzmán, S. (1999). *Matemática II. COU*. Madrid: España: Anaya.

Estepa, A. (1995). Algunas consideraciones sobre la enseñanza de la asociación estadística. *Uno. Revista de Didáctica de las Matemáticas* 5, 69-79.

Estepa, A. y Sánchez Cobo, F. (2001) Correlación y regresión en los primeros cursos universitarios. En *Actas de las Jornadas Europeas de Estadística: la enseñanza y la difusión de la estadística*. Obtenido de la página electrónica [www.caib.es/ibae/esdeveniment/jornadas\\_10\\_01](http://www.caib.es/ibae/esdeveniment/jornadas_10_01)

Fernández, F. y Monroy, O. (1995) Experiencias en el manejo del coeficiente de correlación de Pearson en un curso de estadística. En P. Gómez, C. Carulla, M. Castro, F. Fernández, C. Gómez, V. Mesa, P. Perry & P. Valero (Eds.), *Aportes de 'Una empresa docente' a la IX CIAEM* (pp. 107-117). Bogotá, Colombia: Una empresa docente. Obtenido de la página de internet <http://ued.uniandes.edu.co>.

Fernández, F., Monroy, O. y Rodríguez, L. (1998). *Diseño, desarrollo y evaluación de situaciones problemáticas de estadística*. Bogotá, Colombia: Una empresa docente.

Fones , M. A.(1998). *Matemática 2 Polimodal*. Buenos Aires, Argentina: Kapeluz.

Inhelder, B. y Piaget, J. (1955) *De la logique de l'enfant à la logique de l'adolescent*. Paris, France: Presses Universitaires de France

Rico, L. (1997) Consideraciones sobre el currículo de matemática para la educación secundaria. En Luis Rico (Coord.), *La educación matemática en la enseñanza secundaria* (pp. 15-38). Barcelona, España: ICE-Horsori.

Sánchez Cobo, F.; Estepa, A. y Batanero, C. (2000) Un estudio experimental de la estimación de la correlación a partir de diferentes representaciones. *Enseñanza de las ciencias* 18 (2), 294-310..

Turner, S. (2000) *Matemática 4*. Buenos Aires, Argentina: AZ Editora.



- **Andrea Lina Lavallo**

Departamento de Estadística, Facultad de Economía y Administración  
Universidad Nacional del Comahue

E-mail: alavalle@uncoma.edu.ar

- **Elda Beatriz Micheli**

Departamento de Estadística, Facultad de Economía y Administración  
Universidad Nacional del Comahue

E-mail: elda@uncoma.edu.ar

- **Natalia Rubio**

Departamento de Estadística, Facultad de Economía y Administración  
Universidad Nacional del Comahue

E-mail: rubionatalia@hotmail.com