SUMA 22
junio 1996, pp. 71-78

La calculadora gráfica en correlación y regresión

Luis M. Botella López

IDEAS Y RECURSOS

Se pretende en el presente artículo mostrar el desarrollo del bloque temático de «Correlación y regresión» tal y como se ha llevado a cabo en un grupo de 1.° del nuevo bachillerato, modalidad de Ciencias Sociales, mediante la presentación de algunas de las actividades tratadas. Podemos destacar dos principios básicos de la metodología empleada:

- Las clases se basan de una forma sistemática en la *Resolución de Problemas.*
- Se utiliza, de modo habitual, la calculadora gráfica (en nuestro caso la TI-82) como herramienta de trabajo.

Este bloque temático nos brinda la ocasión de desarrollar y relacionar entre sí una gran cantidad de los conceptos y procedimientos constitutivos del currículum de los nuevos bachilleratos, tanto en la modalidad de Ciencias Sociales como en la de Ciencias de la Naturaleza y en la de Tecnología, al interrelacionar modelos funcionales teóricos con estudios experimentales, análisis con estadística. Disponer en la actualidad de una herramienta como la calculadora gráfica enriquece sobremanera su desarrollo, al permitir abordar el estudio de regresiones no lineales y obviar el farragoso trabajo de realización de cálculos repetitivos; en definitiva, permite dar prioridad al razonamiento sobre el cálculo.

Tras un primer trimestre dedicado a tratar problemas conducentes a asimilar el concepto de relación funcional entre dos magnitudes, y a modelizar las relaciones encontradas, supongo que alumnos y alumnas son capaces de reconocer a partir de una gráfica los tipos de funciones estudiados:

- Lineales, haciendo hincapié en la proporcionalidad directa.
- Proporcionalidad inversa.
- Polinómicas.

Es presciso mencionar que los modelos de función inversa, exponencial y logarítmica fueron estudiados cuando surgió una regresión de uno de estos tipos (en concreto exponencial) durante la resolución de una de las actividades propuestas en este bloque.

El bloque temático se dividió en tres partes (a efecto de proponer actividades con una cierta afinidad):

- Reconocer funciones.
- Relación funcional experimental.
- Relación estadística.

Reconocer funciones

Identificar gráficas

Se trata de reconocer una gráfica proyectada a partir de la calculadora gráfica, e intentar dar su ecuación con el mínimo número de preguntas. Se proyecta:

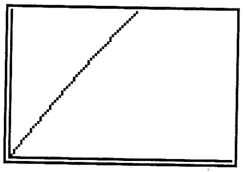

Figura 1

$y = x\,(x\text{-}10)^2$, con escala definida por los parámetros:

$X_{min} = 0$ $X_{max} = 1$ $X_{scl} = 0$

$Y_{min} = 0$ $Y_{max} = 50$ $Y_{scl} = 0$

Creen identificar la gráfica con una recta y piden un valor (x = 1, y = 81), dando como ecuación y = 81x. La representamos gráficamente sobre la anterior[1] con la consiguiente sorpresa (figura 2a). El diálogo que sigue es algo así:

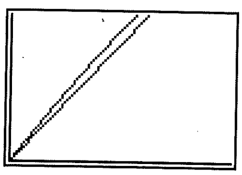

Figura 2a

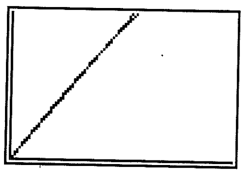

Figura 2b

Alumnos: O sea, que no era una recta

Profesor: Bueno, a lo mejor os habéis confundido con la ecuación.

A: No, porque al pasar por el origen la imagen de 1 nos daría la pendiente.

P: ¿Qué se puede hacer para asegurarse?

A: Veamos en la tabla las imágenes de 0 y 2. (x = 0, y = 0 ; x = 2, y = 128). *Seguro que no lo es, porque las diferencias no son constantes.*

P: De todos modos, ¿por qué no probáis a cambiar la ecuación de la recta?

A: Por ejemplo y = 92x (se dibuja y prácticamente coincide con la gráfica inicial; figura 2b). *Pero no es. Nos hace falta otro valor* (x = 3, y = 147). *Las segundas diferencias tampoco son constantes, no es de segundo grado. Otro valor* (x = 4, y = 144); *¡ya está!, las terceras diferencias sí son constantes. Es de tercer grado... Para tener una idea de la ecuación nos haría falta encontrar las otras raíces... siguiendo con las diferencias terceras, creo que es 10. A ver la imagen* (x = 10, y = 0).

Se trata de reconocer una gráfica proyectada a partir de la calculadora gráfica, e intentar dar su ecuación con el mínimo número de preguntas.

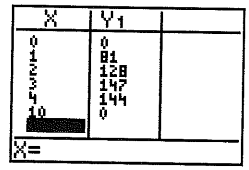

Figura 3

A: Vamos a cambiar la escala para ver la gráfica hasta x = 10.

(Se apaga el retroproyector para hacer $X_{max} = 10$). Ver figura 4a.

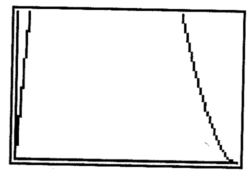

Figura 4a

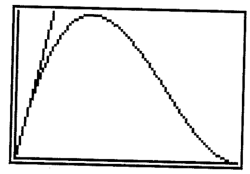

Figura 4b

¡La raíz es doble! Hay que asegurarse; cambio de escala con $X_{max} = 15$, $Y_{max} = ...150$ (tras revisar los valores de la tabla en sus notas, ver figura 3, se obtiene la gráfica de la figura 4b). *¡Sí, es doble! Ya tenemos la ecuación,*

$$y = x(x\text{-}10)^2.$$

Se aprovecha para mostrar el parecido de ambas gráficas (cúbica y recta) en el intervalo [0, 1] y comentar que si deseáramos estudiar la función sólo en ese intervalo, quizá sería más útil tomar una recta a pesar del error cometido. Sobre todo si se tratase de datos que ya tienen algún error. También tenemos ocasión para hablar de los términos «interpolar» y «extrapolar», y del riesgo de hacer predicciones a partir de una tabla de datos para valores muy alejados de los conocidos.

1 Para ello, se apaga el retroproyector, con el fin de mostrar tan sólo las gráficas y que no vean la ecuación buscada hasta que no se considere oportuno.

Buscar gráficas

I. Se visualizan los puntos (40, 60), (60, 100), (65, 110)

A: ¿Cómo quieres que digamos algo con eso?

P: Ahora se trata de que digáis cuál sería la función cuya gráfica pasa por esos puntos.

A: Como tenemos tantos datos... Es una recta, eso sí. ¿Cuáles son los puntos?

P: ¿Cuánta información necesitáis?

A: Las coordenadas de dos puntos.

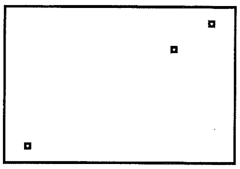

Figura 5

Se mueve el cursor (figura 6a) por dos de los puntos: (x = 40, y = 60), (x = 60, y = 100). Calculan...

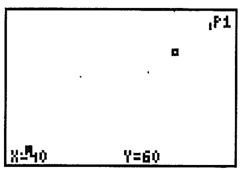

Figura 6a

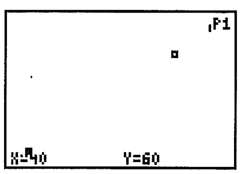

Figura 6b

La y aumenta 40 unidades cuando la x aumenta 20. Así que la pendiente es 2; tenemos x = 40, para ir hasta 0 hay que retroceder 40 unidades, luego la y tiene que disminuir 2·40 = 80, y valía 60. Así que la imagen de 0 es –20. La ecuación es y = 2x – 20.

Se dibuja y, claro, pasa por los puntos.

II. Se visualiza, del mismo modo, la tabla de valores:

50	48	47	44	41	36	35	31	28	26	25	24
49	50	54	57	61	64	59	63	63	61	64	70

Se propone una serie de actividades en que hay que encontrar «la mejor» aproximación a la relación funcional existente entre dos variables, para lo que conocemos una tabla de valores obtenida experimentalmente.

2 Término evidentemente subjetivo en estos momentos.

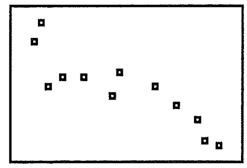

Figura 7a

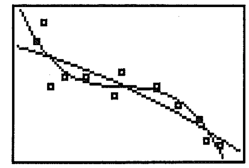

Figura 7b

A: ¡Hala! ¡Cualquiera encuentra una función que pase por esos puntos!

P: ¿De qué tiene eso aspecto?

Diversidad de opiniones.

Pues una recta no.

Puede que de una función de tercer grado.

Yo dibujaría una parábola hacia abajo...

Se puede dar pie a discutir, utilizar las posibilidades de la calculadora para dibujar las mejores aproximaciones para cada una de las ideas (figura 7b), etc. En cualquier caso, se deja el campo abierto para el estudio que viene después.

Relación funcional experimental

Se propone una serie de actividades en que hay que encontrar «la mejor[2]» aproximación a la relación funcional existente entre dos variables, para lo que conocemos una tabla de valores obtenida experimentalmente.

Se trata de que identifiquen la función buscada a partir de la nube de puntos. Una vez hecho, que cada uno construya la función que mejor se adapte a los puntos, para luego discutir los criterios seguidos y concluir cuál parece «la mejor aproximación».

Se utiliza la calculadora para corroborar o refutar hipótesis de forma intuitiva, representando gráficamente la nube de puntos y superponiendo las gráficas de las funciones propuestas.

Aleaciones

El punto de fusión de una aleación depende de las proporciones en que intervienen cada uno de sus componentes. Para dos componentes A y B, se obtienen los siguientes datos:

Proporción de A	0,1	0,3	0,5	0,7	0,9
Temperatura de fusión (°C)	720	580	425	485	555

¿Cuál estimas que es la temperatura de fusión si el componente A aparece al 55%? ¿Cuál crees que será la temperatura de fusión del componente A?

De entrada, identifican los puntos con una parábola. (Caras de desánimo al pensar en el sistema de ecuaciones que les espera; ¡y no pasa por el origen!).

Alguien sugiere un modo de responder a las preguntas: 0,55 es «la mitad de la mitad de la distancia entre 0,5 y 0,7». ¿Por qué no calcular el valor que está a la cuarta parte de 425 a 485? Eso digo yo ¿por qué no? Han hecho una interpolación lineal; a continuación inventan algo parecido para el 1 a partir de los dos últimos datos (extrapolación lineal).

En cualquier caso, es el momento de mostrar cómo la calculadora se hace cargo de lo rutinario y farragoso:

Tras introducir los datos, se define el gráfico deseado, y los vemos en pantalla.

La máquina se hace cargo de los cálculos. Primero definiremos las listas que hacen el papel de variables independiente y dependiente y, a partir de ahí, tenemos todo un repertorio de curvas de regresión. En nuestro caso, elegimos una regresión cuadrática. Ya tenemos calculados los parámetros de «la mejor» aproximación de segundo grado.

Podemos superponer la función así obtenida a la nube de puntos y pasando a la tabla de valores asociada a dicha función damos con los valores deseados.

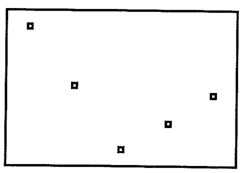

Figura 8a

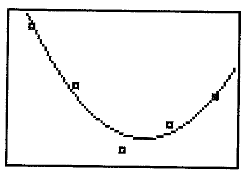

Figura 8b

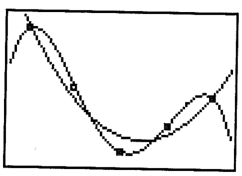

Figura 8c

Después, al haber 5 puntos (condiciones), alguien habla de un polinomio de 4.° grado. Se obtiene con la calculadora y se superpone la gráfica.

Conclusión: No siempre la mejor solución (matemáticamente hablando lo es, puesto que esta gráfica pasa por los cinco puntos) es la más adecuada.

Neumáticos

La profundidad de la huella de la banda de rodadura de un neumático tiene una importante influencia en la conducción. Y no sólo en la evacuación del agua cuando llueve, sino también en la distancia de frenado. La tabla nos da información al respecto:

Profundidad del perfil en mm	8	4	3	2	1
Distancia de frenado en m (a 100 km/h)	70	82	87	97	118

...al permitir la utilización de la calculadora una comprobación inmediata de conjeturas, permiten en algunos minutos toda una serie de reflexiones sobre la influencia de los parámetros en un modelo funcional.

¿Cuál estimas que sería la distancia de frenado para una profundidad de 6 mm?

El límite mínimo legal para la profundidad del dibujo es de 1,6 mm. ¿Cuál será la distancia de frenado?

Hay quien asocia la nube de puntos al modelo de función hiperbólica, con lo que tenemos la ocasión de escuchar razonamientos del siguiente tipo mientras trabajan en grupos con sus calculadoras :

La función es del tipo

$$y = \frac{a}{x} + b$$

La asíntota horizontal debe estar por el 68, así que debe ser b = 68 ; voy a probar con a = 1 y b = 68.

¡Qué va! La a es mucho más grande. Prueba con a=50.

Eso está mejor, pero b es menor; vamos a probar con b = 65.

Pues b más pequeña y a más grande; a = 65 y b = 60.

Todavía no; mira a ver si a = 70.

No, mejor a = 75.

Finalmente obtenemos algo similar a la figura 9. Merece la pena destacar los conocimientos que se ponen en juego en tales discusiones y que, al permitir la utilización de la calculadora una comprobación inmediata de conjeturas, permiten en algunos minutos toda una serie de reflexiones sobre la influencia de los parámetros en un modelo funcional.

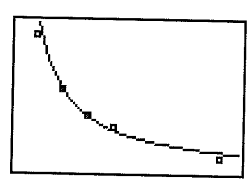

Figura 9

En otros casos, se utiliza la calculadora para ensayar cada una de las opciones disponibles hasta encontrar la gráfica que más se ajusta a los puntos para, a partir de la tabla de valores que corresponda a la función, responder a las preguntas planteadas.

Vemos las gráficas que se obtiene para cada tipo de línea de regresión: tenemos la nube de puntos y las regresiones lineal, cuadrática, cúbica, cuártica y potencial (figura 10).

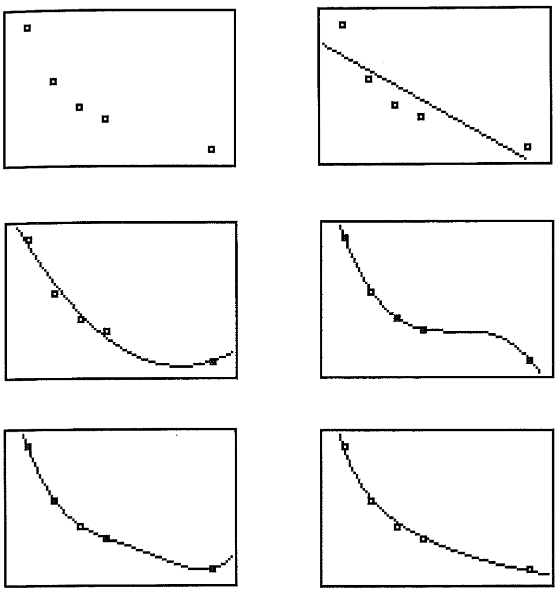

Figura 10

Se plantea la discusión acerca de la idoneidad o no de cada una de las aproximaciones. Ahora se puede responder a las cuestiones planteadas en el enunciado a partir de la tabla de valores:

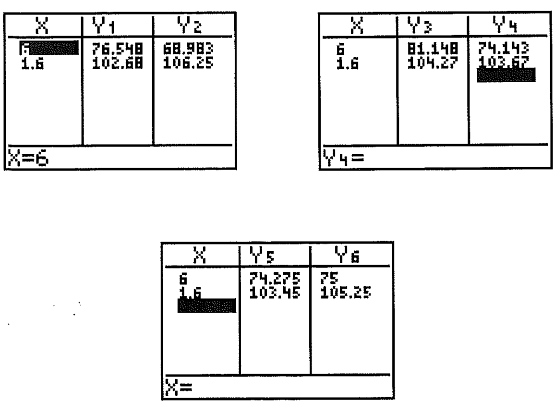

Figura 11

Planetas

El siguiente cuadro da para cada planeta del sistema solar su distancia media al Sol, considerando como 10 la distancia de la Tierra al Sol, así como la duración de un giro completo alrededor del mismo.

Planeta	Número (n)	Distancia (d_n)	Tiempo de revolución (T_n)
Mercurio	1	3,87	88 días
Venus	2	7,23	224,7 días
Tierra	3	10	1 año
Marte	4	15,24	1,88 años
Asteroides	5	29,09	
Júpiter	6	52,03	11,9 años
Saturno	7	95,46	29,5 años
Urano	8	192	84 años
Neptuno	9	300,9	164,8 años
Plutón	10	395	247,7 años

a) Representa gráficamente la relación entre el número del planeta y su distancia al Sol (n y d_n).

b) ¿Qué tipo de relación te parece que puede existir entre ambas variables? Busca un modelo funcional apropiado.

c) ¿A qué distancia sería de esperar que se encontrase un eventual 11.° planeta?

d) Representa gráficamente la relación entre la distancia al Sol y el período de rotación (d_n y T_n).

e) ¿Qué relación funcional encuentras entre ambas? Si lo consigues, acabarás de deducir la tercera ley de Kepler. Búscala en alguna enciclopedia y compara lo que allí pone con lo que tú has encontrado.

Al ensayar, resulta que el mejor ajuste se obtiene con una regresión exponencial para la relación n ↔ d_n. Ello supone un primer contacto con el modelo de función exponencial.

Se explora la gráfica con el cursor o bien se utiliza la tabla de valores para responder a la cuestión b).

Para la relación d_n ↔ T_n se obtiene, aproximadamente, $y = 0{,}32x^{1,5}$, que se deberá relacionar con la tercera ley de Kepler tras recordar o introducir la idea de exponente fraccionario. Las gráficas son:

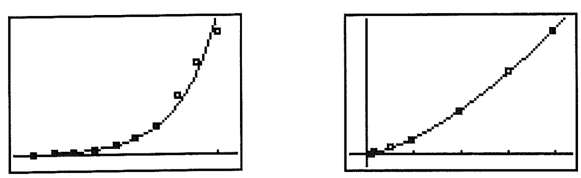

Figura 12

Elementos radiactivos

Los elementos radiactivos emiten partículas de distintos tipos y van perdiendo masa con una rapidez que varía de unos a otros. Un experimentador observa uno de estos elementos en días sucesivos y anotando su masa obtiene la siguiente tabla:

Tiempo (días)	0	1	2	3	4	5	6	7	...
Masa (gramos)	1	0,95	0,90	0,86	0,82	0,78	0,74	0,68	...

Busca un modelo de función que se ajuste a esta tabla.

El periodo de semidesintegración de un elemento radiactivo es el tiempo que tarda una determinada masa del mismo en reducirse a la mitad. ¿Cuál es el periodo de semidesintegración del elemento estudiado?

Teniendo en cuenta esto, quizá el modelo de función que has dado anteriormente deba ser reconsiderado. Hay que tener presente que el período de semidesintegración es fijo; es decir, que cuando vuelva a transcurrir otra vez el mismo tiempo, la masa se volverá a convertir en la mitad y, desde luego, la masa nunca llegará a ser negativa (¿ocurría esto con el modelo que diste antes?).

Busca otro modelo funcional que se ajuste mejor a la situación. ¿Cuándo será la masa menor que 0,1 miligramos? ¿Llegará alguna vez a desaparecer por completo el elemento?

Quizá sea interesante no dar toda la información con el enunciado, sino en dos partes. Al utilizar la calculadora para obtener una línea de regresión adecuada a la nube de puntos, y si se ensaya siguiendo el orden que aparece en la calculadora, es probable que los alumnos y alumnas se conformen con una regresión lineal que, realmente, da una muy buena aproximación (figura 13a).

Se puede ahora preguntar si parece lógico este modelo, que inevitablemente conducirá a una masa negativa tras transcurrir un cierto tiempo (se puede sugerir que recorran la gráfica con el cursor o que visualicen una tabla de valores). Luego han de buscar otro modelo.

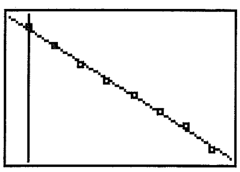

Figura 13a

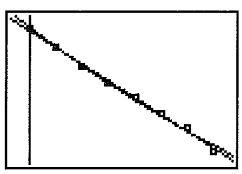

Figura 13b

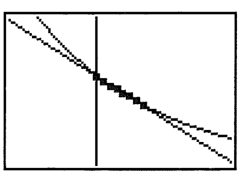

Figura 13c

Al utilizar la calculadora para obtener una línea de regresión adecuada a la nube de puntos, y si se ensaya siguiendo el orden que aparece en la calculadora, es probable que los alumnos y alumnas se conformen con una regresión lineal que, realmente, da una muy buena aproximación.

Al dar la siguiente información sobre el período de semidesintegración se sugiere el modelo de función exponencial (segundo contacto con este modelo), que se ajusta mejor al problema: realizamos un ajuste exponencial y la superponemos a las anteriores. Con un zoom vemos el comportamiento distinto de ambos modelos con el paso del tiempo (figuras 13b y 13c).

Con un zoom o modificando los parámetros de la ventana gráfica podemos dar respuesta al resto de cuestiones planteadas (figura 14), aunque también es posible utilizar las tablas de valores.

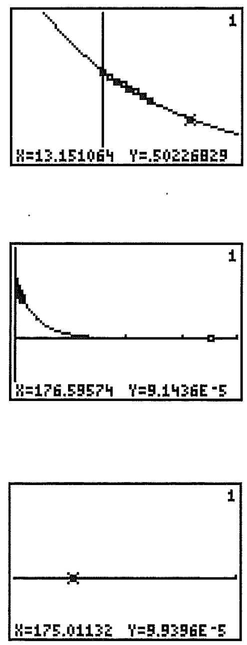

Figura 14

Relación estadística

Equipos de fútbol

En la siguiente tabla tenemos la clasificación de la 1ª división de fútbol el día 10-5-95, junto con los goles marcados por cada equipo y el número de partidos empatados:

	Goles a favor	Partidos empatados
1. Real Madrid	69	9
2 Deportivo	50	10
3. Zaragoza	48	5
4. Betis	38	14
5. Barcelona	51	8
6 Sevilla	45	9
7. Español	41	12
8. At. Bilbao	33	10
9. Oviedo	38	13
10. Valencia	43	10
11. Tenerife	48	7
12. Celta	28	12
13. R. Sociedad	42	13
14. Compostela	35	11
15. At. Madrid	45	7
16. Rácing	35	7
17. Albacete	35	13
18. Sp. Gijón	37	12
19. Valladolid	20	9
20. Logroñés	13	9

a) *¿Qué equipos han marcado pocos goles? ¿Y muchos?*

b) *¿Podrías decir qué lugar debería ocupar en la clasificación un equipo que hubiera marcado 30 goles?*

c) *¿Y un equipo que hubiese empatado 16 partidos?*

Una vez introducidos los datos en tres listas, L_1 a L_3, la calculadora nos da directamente los valores necesarios. El apartado a) se refiere al intervalo

$$\left[\bar{y} - \sigma_y , \bar{y} + \sigma_y\right]$$

obteniendo los parámetros necesarios directamente a partir de los cálculos para dos variables:

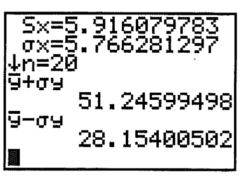

Figura 15

Se puede considerar como «normal» el haber marcado de 29 a 51 goles.

...una vez más la calculadora trabaja por nosotros: activando un gráfico estadístico, calcula la línea de regresión que se copia en la pantalla.

Para el apartado b), una vez más la calculadora trabaja por nosotros: activando un gráfico estadístico, calcula la línea de regresión que se copia en la pantalla. Obtenemos el gráfico adjunto; entramos en la tabla y probamos distintos valores de X hasta obtener el valor más aproximado a 30 para el número de goles (figura 16).

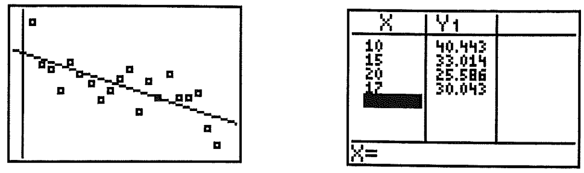

Figura 16

En lo que se refiere al apartado c), hemos de cambiar las definiciones de gráfico para estudiar ahora la relación entre L_1 y L_3.

Al ver el gráfico (figura 17) podemos deducir una débil relación entre las variables, lo que se confirma al calcular la línea de regresión, pues se obtiene como coeficiente de correlación r = 0,086, con lo que no merece la pena proseguir, pues la estimación no es fiable.

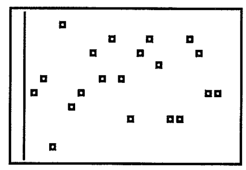

Figura 17

Vehículos y víctimas

El parque de vehículos y el número de víctimas mortales en carretera en España ha evolucionado en los últimos años de la siguiente forma (datos obtenidos de Anuario El País, 1995):

Año	1986	1987	1988	1989	1990	1991	1992	1993
Miles de vehículos	12.284	13.068	13.881	14.870	15.696	16.528	17.347	17.809
Victimas mortales	4.065	5.041	5.419	6.095	5.936	5.744	5.088	4.735

a) *¿Podrías dar una estimación sobre el número de vehículos que se espera en 1994 y 1995?*

b) *¿Y sobre el número esperado de víctimas mortales en estos dos años?*

c) *Quizá se podría pensar que el aumento del número de vehículos debe traer consigo un aumento del número de víctimas mortales. Cita algunas causas por las que creas que no ha sido así en estos ocho años.*

La gráfica definida para responder a la primera cuestión (figura 18a) sugiere una relación de tipo lineal bastante fuerte, obteniendo la respuesta en la tabla de valores (figura 18b):

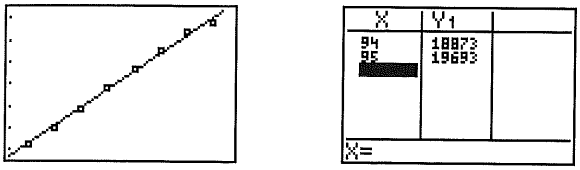

Figura 18

Para la segunda, la gráfica de la figura sugiere una relación no lineal, quizá cuadrática, de la que se obtienen las respuestas de la tabla de la figura 19:

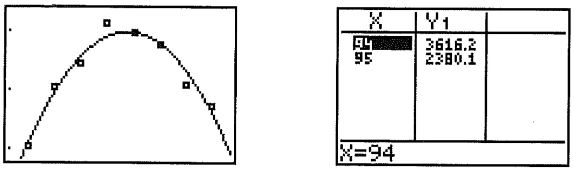

Figura 19

Una conclusión inesperada

Una vez finalizado el bloque, se propuso a alumnas y alumnos algunos problemas que tuvieran relación con todo lo estudiado a lo largo de los dos trimestres del curso. Uno de ellos decía así:

Un intermediario ha comprado el 30 de septiembre 200 kg de uva para venderlos en Navidad. Le han costado a 40 pta el kilo y sabe que cada día que pase, su precio aumentará en 1 pta/kg, pero pierde 1 kg de uva.

a) *Encuentra una función que proporcione el importe de la venta de las uvas (si el intermediario vendiese todos los kilos disponibles) cuando hayan transcurrido x días y represéntala gráficamente.*

b) *¿Cuándo le interesa vender? ¿Qué ganancia puede obtener como máximo?*

Uno de los grupos desarrolló el problema del siguiente modo:

1. Elaboración de una tabla de valores construida a partir del enunciado; introducen los datos en la calculadora y visualizan la nube de puntos

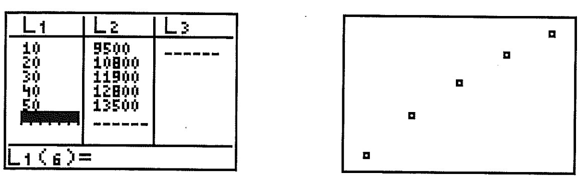

Figura 20

2. Intuyen que se trata de una función polinómica de 2.° grado (¿o quizá prueban antes un ajuste lineal y encuentran la coincidencia al segundo intento?), realizan el ajuste correspondiente y superponen la gráfica a la nube de puntos:

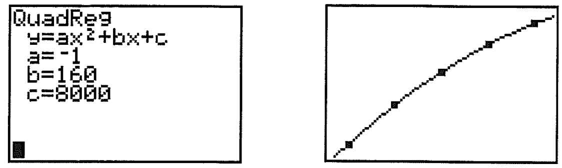

Figura 21

3. Para asegurarse, añaden un punto a la nube, separado de los otros, y dibujan el resultado:

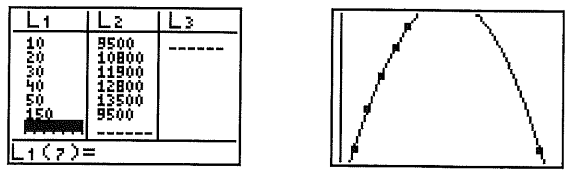

Figura 22

4. Cambian los valores que definen la ventana gráfica:

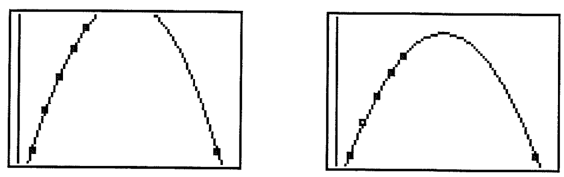

Figura 23

5. Ya seguros de lo que hacen, dan como expresión algebraica la así obtenida, y para el máximo exploran la gráfica con el cursor y ajustan con el zoom. Se aseguran a partir de la tabla de valores.

Luis M. Botella
IB Francisco Pacheco
Alicante
Societat d'Educació
Matemàtica
Al-Khwarizmi

Bibliografía

GUZMÁN, M. y J. COLERA (1989): *COU. Matemáticas II*, Anaya, Madrid.

ENGEL, A. (1990): *Les certitudes du hasard*, Aleas Editeur, Lyon.