

Modelos lineales de ajuste de datos

**Antonio R. Quesada
Nicolás Rosillo Fernández**

LOS AVANCES y la masificación de ordenadores personales y otros productos electrónicos, así como la difusión de los medios de comunicación y de internet, están poniendo a nuestra disposición una avalancha, cada vez mayor, de información. Afortunadamente, tanto los nuevos programas de estadística para ordenadores como las calculadoras modernas nos permiten almacenar y analizar datos con un mínimo de esfuerzo y de tiempo. Esta situación, que está creciendo exponencialmente, sugiere la necesidad de equipar a nuestros estudiantes de secundaria con las herramientas y el conocimiento necesario para llevar a cabo análisis de datos. No es de extrañar que el análisis de datos aparezca en las últimas recomendaciones curriculares (NCTM, 1989), y que haya empezado a aparecer en los libros de texto (Core-Plus Mathematics Project, 1997; Demana & Waits, 1997; The North Carolina School of Science and Mathematics, 1992), así como en manuales de reciente publicación (Morgan, 1997).

En este artículo se presentan y contrastan los dos modelos lineales que están disponibles en las calculadoras Texas Instruments TI-82 y TI-83, al mismo tiempo que se demuestra la facilidad con que algunos de los cálculos tradicionales relacionados pueden llevarse a cabo cuando, por razones de tipo pedagógico, se estime conveniente.

Deliberadamente, en deferencia al lector poco familiarizado con estas calculadoras, se han usado una profusión de figuras para describir los distintos procesos que se llevan a cabo, de forma que sea fácil reproducirlos.

La recta de mínimos cuadrados

En primer lugar se considera un ejemplo con el que se ilustra el cálculo de la recta de regresión tradicional que,

En este artículo se presentan y contrastan, de acuerdo tanto a la suma de los cuadrados de los residuos como a la existencia de puntos atípicos, los dos modelos lineales que están disponibles en las calculadoras Texas Instruments TI-82 y TI-83. Al mismo tiempo se demuestra la facilidad con que algunos de los cálculos tradicionales relacionados pueden llevarse a cabo cuando, por razones de tipo pedagógico, se estime conveniente.

como es bien sabido, se obtiene de forma que la suma de los cuadrados de los residuos (distancia vertical de un punto dado a la recta) sea mínima.

Ejemplo 1. La tabla que sigue contiene el valor esperado de vida para los ciudadanos estadounidenses en el año 1989. Encuentre un modelo que describa este conjunto de datos y úselo para estimar la esperanza de vida de dos personas con edades de 85 y 43 años respectivamente.

Edad	10	20	30	40	50	60	70	80
Esperanza de vida	66	56	47	37	29	20	14	8

Tabla 1. Centro Nacional de Estadísticas de la Salud de EEUU

Una vez que se entran los datos de la tabla en dos listas, digamos L_1 y L_2 en este caso (figura 1.a), se procede a obtener el diagrama de dispersión seleccionando, como ilustra la figura 1.b, el tipo de grafo y especificando las listas que contienen los datos.

L1	L2	L3	1
66	66		----
56	56		
47	47		
37	37		
29	29		
20	20		
14	14		

L1()=10

Figura 1.a

```

2ND F1 Plot2 Plot3
On Off
Type: [ ] [ ] [ ]
Xlist: L1
Ylist: L2
Mark: [ ] +
    
```

Figura 1.b

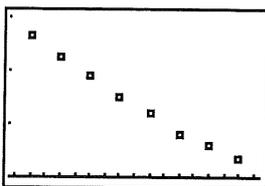


Figura 1.c

```

WINDOW
Xmin=3
Ymin=-1.86
Xmax=87
Ymax=75.86
Xscl=5
Yscl=25
Xres=1
    
```

Figura 1.d

En este caso se ha dejado que la calculadora determine las dimensiones de la ventana (figura 1.d) usando *ZoomStat*. Ya que el diagrama parece ser lineal (figura 1.c), se selecciona la recta de mínimos cuadrados (figura 2.a) y en la pantalla base se especifican las listas que contienen los datos, así como la función en la que se quiere almacenar la ecuación (figura 2.b), aunque esto último es opcional (disponible sólo en la TI-83). La calculadora muestra los coeficientes de la ecuación de la recta en la pantalla principal (figura 2.c) a la vez que escribe la ecuación en la función Y_1 elegida (figura 2.d).

A fin de obtener el coeficiente de correlación r y su cuadrado r^2 junto a los coeficientes de la recta, se debe seleccionar y ejecutar la instrucción *DiagnosticOn* del CATALOG (disponible sólo en la TI-83). La figura 3.a muestra

que la gráfica de la recta se ajusta casi perfectamente a los puntos, lo que confirma el coeficiente de correlación obtenido de casi -1 . Una vez obtenido el modelo lineal, se puede sencillamente evaluar la función Y_1 usando la tabla (figuras 3.b y 3.c) o la pantalla principal (figura 3.d) para estimar la esperanza de vida correspondiente a las edades dadas de 43 y 85 años. Sería interesante discutir con los estudiantes los valores esperados para personas de 86 o de 90 años.

```

EDIT [ ] TESTS
1:1-Var Stats
2:2-Var Stats
3:Med-Med
4:LinReg(ax+b)
5:QuadReg
6:CubicReg
7:QuartReg
    
```

Figura 2.a

```

LinReg(ax+b) L1,
L2, Y1
    
```

Figura 2.b

```

LinReg
y=ax+b
a=-.8392857143
b=72.39285714
r^2=.9928208877
r=-.9964039782
    
```

Figura 2.c

```

2ND F1 Plot2 Plot3
Y1=-.8392857142
8571X+72.3928571
42857
Y2=
Y3=
Y4=
Y5=
    
```

Figura 2.d

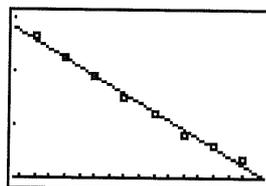


Figura 3.a

```

TABLE SETUP
TblStart=0
ΔTbl=1
Indent: Auto
Depend: Horiz Hsk
    
```

Figura 3.b

X	Y1
85	1.0536
43	36.304

X=

Figura 3.c

```

Y1(85)
1.053571429
Y1(43)
36.30357143
    
```

Figura 3.d

Tradicionalmente, dado el conjunto de puntos $\{(x_i, y_i) : 1 \leq i \leq n\}$, la recta de mínimos cuadrados $y = ax + b$ que los modela se calcula usando las expresiones:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \quad [1]$$

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

En la figura 4.a se han usado estas fórmulas para obtener los coeficientes de la recta de regresión. Las figuras 4.b y 4.c ilustran los mismo cálculos usando el tipo de datos lista y las operaciones asociadas.

```
(n(Σxy)-ΣxΣy)/(n
Σx²-(Σx)²)+A
-.8392857143
n¹(Σy-AΣx)
72.39285714
```

Figura 4.a

```
(dim(L1)sum(L1L2)
)-sum(L1)sum(L2)
)/(dim(L1)sum(L1
²)-(sum(L1))²)+A
-.8392857143
```

Figura 4.b

```
(dim(L1)-1)(sum(L
z)-A*sum(L1))
72.39285714
```

Figura 4.c

La variedad de representaciones disponibles en las nuevas calculadoras permiten abordar un problema usando enfoques completamente distintos. Así, por ejemplo, a continuación se considera cómo obtener la recta de regresión usando sistemas de ecuaciones lineales. Ya que todo par ordenado de la tabla debe satisfacer la ecuación de la recta $y = ax + b$ que se busca, sustituyendo se obtiene el sistema lineal que sigue y que se ha expresado en forma matricial.

$$\begin{cases} 10a + b = 66 \\ 20a + b = 56 \\ M \\ 80a + b = 8 \end{cases} \quad AX = B, \text{ donde } A = \begin{bmatrix} 10 & 1 \\ 20 & 1 \\ M & M \\ 80 & 1 \end{bmatrix}, X = \begin{bmatrix} a \\ b \end{bmatrix} \text{ y } B = \begin{bmatrix} 66 \\ 56 \\ M \\ 8 \end{bmatrix}$$

Multiplicando por la izquierda ambos lados de la ecuación matricial por A^T se consigue que la matriz $A^T A$, de los coeficientes de X , sea cuadrada, lo que permite resolver por X usando la inversa, esto es,

$$A^T A X = A^T B \text{ de donde } X = (A^T A)^{-1} A^T B.$$

Las figuras 5.a y 5.b muestran las matrices A y B mientras que la figura 5.c contiene los cálculos necesarios y la matriz resultante que de nuevo confirma los resultados anteriores.

```
MATRIX[A] 8 x2
[ 10 1
[ 20 1
[ 30 1
[ 40 1
[ 50 1
[ 60 1
[ 70 1
```

Figura 5.a

```
MATRIX[B] 8 x1
[ 66
[ 56
[ 4
[ 20
[ 20
[ 20
[ 20
[ 14
```

Figura 5.b

```
([A]ᵀ[A])⁻¹[A]ᵀ[B]
[[-.8392857143]
[72.39285714]]
```

Figura 5.c

Así como antes se obtuvieron los coeficientes de la recta de regresión a partir de sus expresiones formales [1], del mismo modo podemos, por razones pedagógicas, estar interesados en obtener el coeficiente de correlación a partir de alguna de las expresiones que lo definen:

$$r = \frac{\sum_i x_i y_i - \left(\sum_i x_i \sum_i y_i \right) / n}{(n-1)S_x S_y} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{(n-1)S_x S_y} = \frac{1}{n-1} \sum_i \frac{x_i - \bar{x}}{S_x} \cdot \frac{y_i - \bar{y}}{S_y} \quad [2]$$

La figura 6.b muestra como calcular r en la pantalla principal usando las variables de la calculadora de acuerdo a la primera y segunda fracción. También se pueden normalizar primero las listas de datos (figura 6.c) y usar las listas resultantes (figura 6.d).

```
XY EQ TEST PTS
1: Σx
2: Σx²
3: Σy
4: Σy²
5: Σxy
```

Figura 6.a

```
(Σxy-(ΣxΣy)/n)/(
(n-1)SxSy)
-.9964039782
sum((L1-Σ)/
)/(n-1)SxSy)
-.9964039782
```

Figura 6.b

```
L2 L3 L4
66 -1.429
56 -1.021
47 -.6124
37 -.2041
29 .20412
20 .61237
14 1.0206
L4=(L2-9)/Sx
```

Figura 6.c

```
sum(L3L4)/(n-1)
-.9964039782
```

Figura 6.d

Antes de presentar el segundo modelo lineal, conviene ponderar el porqué puede ser deseable el uso de otro modelo.

Ejemplo 2. Consideremos la siguiente tabla de valores cuyos datos se han almacenado en las listas L₁ y L₂ (figura 7.a).

x	1	3	4	5	7	8	9	10	11	12	14	16	23
y	7	11	13	15	15	17	16	18	3	18	20	21	4

Tabla 2

La figura 7.b muestra que, con la excepción de dos puntos alejados, el resto de los puntos sugieren una tendencia lineal. Ahora bien, el efecto de estos dos puntos atípicos en la recta de regresión es contundente. Como vemos en la figura 7.c el coeficiente de correlación es casi nulo, a la vez que, como nos muestra la figura 7.d, la gráfica de la recta de regresión está «alejada» de la mayoría de los puntos. No sería razonable, pues, el hacer proyecciones usando este modelo. A continuación, se presenta un modelo lineal alternativo más resistente a un número reducido de valores extremos.

L1	L2	L3	3
1	7		
2	11		
3	13		
4	15		
5	15		
6	17		
7	16		
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			
101			
102			
103			
104			
105			
106			
107			
108			
109			
110			
111			
112			
113			
114			
115			
116			
117			
118			
119			
120			
121			
122			
123			
124			
125			
126			
127			
128			
129			
130			
131			
132			
133			
134			
135			
136			
137			
138			
139			
140			
141			
142			
143			
144			
145			
146			
147			
148			
149			
150			
151			
152			
153			
154			
155			
156			
157			
158			
159			
160			
161			
162			
163			
164			
165			
166			
167			
168			
169			
170			
171			
172			
173			
174			
175			
176			
177			
178			
179			
180			
181			
182			
183			
184			
185			
186			
187			
188			
189			
190			
191			
192			
193			
194			
195			
196			
197			
198			
199			
200			
201			
202			
203			
204			
205			
206			
207			
208			
209			
210			
211			
212			
213			
214			
215			
216			
217			
218			
219			
220			
221			
222			
223			
224			
225			
226			
227			
228			
229			
230			
231			
232			
233			
234			
235			
236			
237			
238			
239			
240			
241			
242			
243			
244			
245			
246			
247			
248			
249			
250			
251			
252			
253			
254			
255			
256			
257			
258			
259			
260			
261			
262			
263			
264			
265			
266			
267			
268			
269			
270			
271			
272			
273			
274			
275			
276			
277			
278			
279			
280			
281			
282			
283			
284			
285			
286			
287			
288			
289			
290			
291			
292			
293			
294			
295			
296			
297			
298			
299			
300			
301			
302			
303			
304			
305			
306			
307			
308			
309			
310			
311			
312			
313			
314			
315			
316			
317			
318			
319			
320			
321			
322			
323			
324			
325			
326			
327			
328			
329			
330			
331			
332			
333			
334			
335			
336			
337			
338			
339			
340			
341			
342			
343			
344			
345			
346			
347			
348			
349			
350			
351			
352			
353			
354			
355			
356			
357			
358			
359			
360			
361			
362			
363			
364			
365			
366			
367			
368			
369			
370			
371			
372			
373			
374			
375			
376			
377			
378			
379			
380			
381			
382			
383			
384			
385			
386			
387			
388			
389			
390			
391			
392			
393			
394			
395			
396			
397			
398			
399			
400			

Figura 7.a

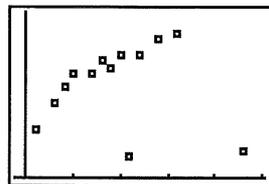


Figura 7.b

```

LinReg
y=ax+b
a=.0113431761
b=13.5849838
r2=1.33824E-4
r=.0115682318
    
```

Figura 7.c

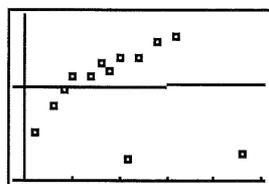


Figura 7.d

La recta med-med

El tipo de situación descrita en el ejemplo anterior lleva a John W. Tukey (1977) a inventar la recta med-med (mediana-mediana) como una recta de ajuste más resistente a los efectos de puntos atípicos, o puntos anormalmente alejados de la mayoría de los puntos del conjunto de datos.

Algoritmo para obtener la recta med-med

Dado un conjunto S de n puntos se procede así:

... lleva a John W. Tukey (1977) a inventar la recta med-med (mediana-mediana) como una recta de ajuste más resistente a los efectos de puntos atípicos, o puntos anormalmente alejados de la mayoría de los puntos del conjunto de datos.

1. Se ordenan los puntos de S en orden creciente de abscisas.
2. Se subdivide el conjunto ordenado de datos en 3 grupos, digamos G₁, G₂ y G₃. Si el número de puntos de S es múltiplo de 3, esto es si n = 3k, entonces cada grupo constará de k puntos. Si n = 3k+1, se deja que el segundo grupo G₂ conste de k+1 puntos. Finalmente si n = 3k+2, se asignan k+1 puntos a los grupos de los extremos. La única excepción a esta regla se hace cuando dos pares tienen la misma abscisa, en cuyo caso se ubican en el mismo grupo.
3. Se procede a calcular para cada grupo G_i, 1 ≤ i ≤ 3, el punto P_i, cuyas coordenadas (x_i, y_i) son, respectivamente, la mediana de las abscisas y la mediana de las ordenadas de los puntos en el grupo.
4. Se obtiene la ecuación de la recta t, de ecuación y = mx + b, que pasa por P₁(x₁, y₁) y P₃(x₃, y₃).
5. Por último, se desplaza la recta t un tercio de la distancia vertical entre el punto P₂(x₂, y₂) y la recta, esto es

$$d = \frac{1}{3} |(mx_2 + b) - y_2|$$

El desplazamiento se hará siempre hacia (x₂, y₂).

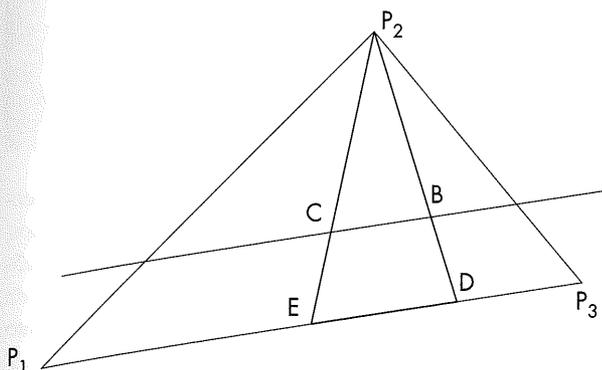
Existe otra formulación alternativa de este quinto paso:

- 5b. Por último, se desplaza la recta t paralelamente a sí misma obligándola a que contenga al punto (a, b), cuyas componentes son

$$a = \frac{x_1 + x_2 + x_3}{3} \quad \text{y} \quad b = \frac{y_1 + y_2 + y_3}{3}$$

Cualquiera de las formulaciones implica la otra. Para ver esto considérese el triángulo que tiene por vértices P₁(x₁, y₁), P₂(x₂, y₂) y P₃(x₃, y₃) como el de la figura. La recta med-med contendrá el lado formado por P₁ y P₃ y ha de desplazarse un tercio de la distancia que separa D de P<

centro del triángulo descrito (punto C de la figura, siendo P_2E mediana), o lo que es lo mismo, por su centro de gravedad.



Los triángulos P_2CB y P_2ED son semejantes, lo que indica que

$$EC = \frac{1}{3}EP_2 \Leftrightarrow DB = \frac{1}{3}DP_2$$

Ejemplo 3. A continuación se aplica el algoritmo descrito para hallar la recta med-med asociada a la distribución del ejemplo 2.

Los subgrupos y puntos correspondientes son:

- $G_1 = \{(1, 7), (3, 11), (4, 13), (5, 15)\}$ $P_1 = (3,5, 12)$
- $G_2 = \{(7, 15), (8, 17), (9, 16), (10, 18), (11, 3)\}$ $P_2 = (9, 16)$
- $G_3 = \{(12, 18), (14, 20), (16, 21), (23, 4)\}$ $P_3 = (15, 19)$

La recta que pasa por los puntos P_1 y P_3 es:

$$y - 19 = \frac{19 - 12}{15 - 3,5}(x - 15) \Rightarrow y = \frac{14}{23}x + \frac{227}{23} \approx 0,609x + 9,870$$

Siguiendo el paso 5, para $x = 9$, se obtiene

$$y = \frac{14 \cdot 9}{23} + \frac{227}{23} = \frac{353}{23} \Rightarrow d = \frac{1}{3} \left| \frac{353}{23} - 16 \right| = 0,217$$

por tanto, si la recta que pasa por P_1 y P_3 tiene como ordenada en el origen 9,870, la recta med-med tendrá $9,870 + 0,217 = 10,087$, con lo que su ecuación será $y = 0,609x + 10,087$

Si seguimos las indicaciones del paso 5. b, la recta med-med será de la forma

$$y = \frac{14}{23}x + n$$

donde n se obtendrá obligando a dicha recta a que pase por el punto (a, b) de coordenadas

$$a = \frac{3,5 + 9 + 15}{3} = \frac{55}{6} \quad \text{y} \quad b = \frac{12 + 16 + 19}{3} = \frac{47}{3}$$

Por tanto

$$\frac{47}{3} = \frac{14}{23} \cdot \frac{55}{6} + n \Rightarrow n = \frac{696}{69} = 10,087$$

lo que indica que la recta med-med tiene de ecuación

$$y = \frac{14}{23}x + \frac{696}{69} \approx 0,609x + 10,087$$

que, como se observa en la figura 8.e, produce un ajuste bastante mejor que la recta de regresión antes hallada. Las figuras 8.a, 8.b, 8.c y 8.d ilustran cómo obtener la recta med-med e incluso los tres puntos P_1, P_2 y P_3 que ayudan a determinarla.

```
EDIT [F2] TESTS
1:1-Var Stats
2:2-Var Stats
3:Med-Med
4:LinReg(ax+b)
5:QuadReg
6:CubicReg
7:QuartReg
```

Figura 8.a

```
Med-Med L1,L2,Y2
█
```

Figura 8.b

```
Med-Med
y=ax+b
a=.6086956522
b=10.08695652
```

Figura 8.c

```
(x1,x2,x3)
(3.5 9 15)
(y1,y2,y3)
(12 16 19)
█
```

Figura 8.d

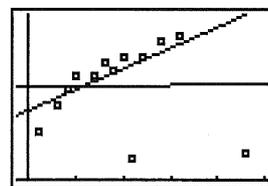


Figura 8.e

El efecto de los puntos patológicos

A la hora de comparar los dos modelos lineales considerados conviene analizar tanto la suma de los cuadrados de los residuos como el efecto de los puntos patológicos en ambos modelos. Para la distribución de la tabla 2, se puede comprobar que la recta de regresión minimiza la suma de los cuadrados de los residuos sólo con introducir en L_3 y L_4 los residuos correspondientes a los dos modelos lineales usados. Ya que en Y_1 se guarda la recta de mínimos cuadrados y en Y_2 la med-med, se definen $L_3 = L_2 - Y_1(L_1)$ y $L_4 = L_2 - Y_2(L_1)$. Debe señalarse que en la TI-83 la lista de residuos se calcula automáticamente con cada modelo lineal y se almacena como la lista LRESID. Por tanto, bastaría definir $L_3 = LRESID$, respectivamente $L_4 = LRESID$, inmediatamente después de calcular cada una de las rectas. Como asegura la teoría y se comprueba en la figura 9.b, la recta de mínimos cuadrados da la menor suma, pero como se ha observado en la figura

8.e, tener la menor suma no garantiza el mejor ajuste. Ahora bien, suprimiendo en L_1 y L_2 las entradas correspondientes a los pares de puntos patológicos (11, 3) y (23, 4), y recalculando la recta de mínimos cuadrados, la nueva recta posee una pendiente y una ordenada en origen mucho más próximas a las dadas por la recta med-med obtenida del conjunto inicial de datos (figura 9.d). Este es el principal valor de la recta med-med: recoger la tendencia de la mayoría de los puntos sin necesidad de remover los puntos patológicos del conjunto de datos.

L2	L3	L4	4
7	-6.5986	-----	
11	-2.819		
13	1.6304		
15	1.40082		
17	2.40082		
19	2.40082		
21	2.40082		
23	2.40082		
25	2.40082		
27	2.40082		
29	2.40082		

$L4 = L2 - Y2(L1)$

Figura 9.a

```
sum(L3^2)
      410.71426
sum(L4^2)
      623.4706994
```

Figura 9.b

```
LinReg(ax+b) L1,
L2, Y3
```

Figura 9.c

```
LinReg
y=ax+b
a=.816872428
b=8.936213992
r^2=.8948630075
r=.9459719909
```

Figura 9.d

La nueva recta de mínimos cuadrados está marcada con trazo grueso en la figura 10.a, y la suma de sus residuos al cuadrado se observa en la figura 10.b.

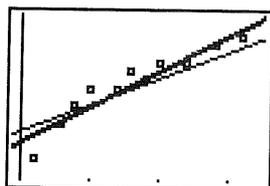


Figura 10.a

```
sum(L5^2)
      17.31893004
```

Figura 10.b

Es un ejercicio sencillo el construir conjuntos de datos para comprobar empíricamente el efecto de los puntos atípicos en los modelos lineales estudiados. Así por ejemplo, podemos pedir a los estudiantes que calculen la recta de mínimos cuadrados y la recta med-med para los tres conjuntos de datos de la tabla 3, y que describan gráficamente y algebraicamente como los puntos atípicos (que en este caso son los que no siguen la relación $y = 2x - 1$) afectan los resultados obtenidos.

x	2	4	6	8	10	12	14	16	18	20
Y_1	3	7	11	15	19	23	27	31	35	23
Y_2	3	7	11	15	38	23	27	31	35	39
Y_3	7	3	11	15	19	23	27	31	35	39

Tabla 3

La elección de un modelo u otro depende de si queremos incluir todos los datos, extremos o no, o si consideramos suficiente el captar la tendencia de la mayoría de los puntos.

Es importante indicar que si bien cuando se encuentran puntos atípicos se deben revisar los datos para asegurarse de que no se trata de un error, estos puntos a menudo pueden denotar alguna característica adicional del modelo que no se ha considerado.

Conclusión

La elección de un modelo u otro depende de si queremos incluir todos los datos, extremos o no, o si consideramos suficiente el captar la tendencia de la mayoría de los puntos. Es claro que si el conjunto de datos es suficientemente grande, la eliminación de los puntos atípicos puede no ser inmediata. En todo caso la decisión, como hemos visto, está entre la recta que minimiza las distancias verticales entre los datos reales y los predichos o la recta menos afectada por los puntos patológicos. En general, la existencia de puntos atípicos cerca del centro del conjunto tiende a afectar la intersección con el eje de ordenadas de la recta de mínimos cuadrados, mientras que cuando estos puntos ocurren cerca de los extremos del conjunto de datos afectan la pendiente de esta recta. De hecho, la influencia de un punto en la pendiente de la recta es directamente proporcional a su distancia al centro del conjunto de datos. En contraste, la recta med-med es relativamente resistente a los efectos de puntos atípicos.

Bibliografía

- CORE-PLUS MATHEMATICS PROJECT (1997): *Contemporary Mathematics In Context, A Unified Approach. Course One*, Janson Publications, Inc., Chicago, IL.
- DEMANA, F. y B. WAITS (1997): *Precalculus*, Addison-Wesley Pub. Company, Reading, Massachussets.
- MORGAN, L. (1997): *Statistics Handbook for the TI-83*, Texas Instruments.
- N.C.T.M. (1989): *Curriculum and Evaluation Standards for School Mathematics*, NCTM, Reston, Virginia.
- THE NORTH CAROLINA SCHOOL OF SCIENCE AND MATHEMATICS (1992): *Contemporary Precalculus Through Applications*, Janson Publ., Inc., Deham, Massachussets.

Antonio R. Quesada
 Department
 of Mathematical Sciences
 The University of Akron
Nicolás Rosillo
 IES Francisco Nieva
 Valdepeñas
 (Ciudad Real)