ESTIMACIÓN DE LA DEMANDA EN PUBLICACIONES PERIÓDICAS CON MÚTIPLES PUNTOS DE VENTA Y STOCKOUTS INFRECUENTES

J.M. Eguzkitza Arrizabalaga.
Departamento de Matemática Aplicada. Universidad del
País Vasco/Euskal Herriko Unibertsitatea.

ABSTRACT
Historical sales and demand are not equivalent when referred to periodicals, since the former are limited by the available stock. When the mean stockout is low and there is not possibility of updating data during the selling period, demand is estimated by means of a maximum-likelihood procedure. The paper also reports another practical approach wich consists of estimating the lack of issues per out-of-stock period and per point-of-sale. In order to compare both methods, real data on sales of a monthly periodical from Bizkaia are used. The results obtained after using the approximate method do practically coincide with those obtained after using the maximum-likelihood procedure. Moreover, computations were much simpler.

RESUMEN
En publicaciones periódicas el historial de ventas no equivale a la demanda, al estar aquéllas limitadas por el inventario disponible. Cuando el nivel promedio de stockouts es bajo y no hay posibilidad de ir actualizando los datos durante el período de ventas, la demanda es estimada mediante un procedimiento de máxima verosimilitud. En este trabajo se propone además otro método aproximado, consistente en efectuar una estimación de los ejemplares faltantes para cada período fuera de stock y para cada punto de venta. Con el fin de comparar ambos métodos se utilizan datos reales de ventas de una publicación mensual en la provincia de Bizkaia. Se comprueba que los resultados obtenidos con el método aproximado son prácticamente coincidentes con los obtenidos por el método de máxima verosimilitud requiriéndose, sin embargo, un esfuerzo de cómputo muchísimo menor.

1. INTRODUCCIÓN

El problema principal que han de enfrentar las distribuidoras de prensa es la optimización del reparto de la tirada de publicaciones periódicas en múltiples puntos de venta. Como las ventas están limitadas por el inventario disponible, el historial de ventas no equivale a la demanda y, en consecuencia, para estimar ésta se parte de una serie de datos censurados por el stock disponible.

La política que, en la mayoría de los casos, siguen las editoras de prensa es fidelizar al cliente colocando en los puntos de venta un número de ejemplares suficiente para que el nivel de servicio sea alto, por lo que el nivel de períodos fuera de stock, y por tanto la tasa de datos censurados, es relativamente bajo (menor del 30%).

Se supone que la demanda es independiente del particular período de tiempo, suposición plausible para muchas publicaciones cuyo patrón de demanda no se ve afectado por estacionalidad, tendencia u otros factores relacionados con el tiempo. Se supone además que los clientes llegan aleatoriamente durante el período de venta, por lo que la demanda sigue una distribución de Poisson, cuyo parámetro λ tratamos de estimar.

Las distribuidoras de publicaciones periódicas se ven en la necesidad de estimar la demanda para muchos puntos de venta y para muchas publicaciones, teniendo en cuenta, además, que esa estimación debe ser hecha con relativa inmediatez cuando se trata de diarios. Se plantea, por tanto, la necesidad de disponer de un método relativamente sencillo para la estimación de λ que no exija grandes requerimientos computacionales.

Con las suposiciones hechas se plantea en primer lugar un procedimiento de máxima verosimilitud (MV), que es una extensión del propuesto por Conrad (1976), pero con distintos niveles de stock para cada período, pues la distribuidora se ve forzada a colocar en un mismo punto de venta diferentes cantidades de ejemplares para cada período, ya que, por razones desconocidas para la distribuidora, la editorial suministra para cada número cantidades globales que varían sensiblemente. Como señalan Lau y Lau (1996), este procedimiento es apropiado cuando las observaciones del lado izquierdo conocido representan un sustancial porcentaje (más del 60%) de la distribución de demanda completa, lo que suele ocurrir, como se ha señalado más arriba, en las distribuidoras de prensa. Como alternativa al método MV se presenta posteriormente un método aproximado (MA) basado en la estimación de la demanda insatisfecha esperada.

Otros métodos de estimación, como el planteado por Lau y Lau (1996), no resultan aplicables en el caso presente pues, entre otras cosas, exigen que los puntos de venta aporten periódicamente datos de ventas durante el período correspondiente, algo prácticamente imposible, pues los quioscos de prensa son ajenos a la empresa distribuidora.

2. ESTIMACIÓN POR MÁXIMA VEROSIMILITUD

Se define:

D_t = Demanda en el período t
S_t = Ventas en el período t
I_t = Stock disponible para el período t

Entonces,

$S_t = min(D_t,I_t)$
$D_t = S_t$ si $S_t < I_t$
$D_t \geq S_t$ si $S_t = I_t$

Los datos de partida, correspondientes a n períodos, son:

Ventas	S_1	S_2	...	S_r	S_{r+1}	S_{r+2}	…	S_n
Stock disponible	I_1	I_2	...	I_r	I_{r+1}	I_{r+2}	…	I_n

Cuando $S_i < I_i$ $(i=1,2,...,r)$ S_i es un verdadero reflejo de la demanda, pero cuando $S_i = I_i$ $(i=r+1,...,n)$ S_i subestima la verdadera demanda. Los datos no siguen necesariamente un orden cronológico, lo que no implica pérdida de generalidad ya que se supone que la demanda es independiente del tiempo.

Considerando que la distribución teórica de demanda es Poisson de parámetro λ, la función de verosimilitud de la muestra será

$$L(\lambda) = \prod_{i=1}^{r}\left(\frac{\lambda^{S_i}e^{-\lambda}}{S_i\,!}\right)\prod_{i=r+1}^{n}\left(\sum_{x=I_i}^{\infty}\frac{\lambda^{x}e^{-\lambda}}{x\,!}\right)$$

es decir,

$$\log L(\lambda) = \log\lambda.\sum_{i=1}^{r} S_i - \lambda r + cte. + \sum_{i=r+1}^{n}\log\sum_{x=I_i}^{\infty}\frac{\lambda^{x}e^{-\lambda}}{x\,!}$$

Derivando respecto de λ e igualando a cero se obtiene la ecuación de estimación:

$$\frac{1}{\lambda}\sum_{i=1}^{r} S_i - r + \sum_{i=r+1}^{n}\frac{\sum_{x=I_i}^{\infty}\left(\frac{x\lambda^{x-1}e^{-\lambda}}{x\,!} - \frac{\lambda^{x}e^{-\lambda}}{x\,!}\right)}{\sum_{x=I_i}^{\infty}\frac{\lambda^{x}e^{-\lambda}}{x\,!}} =$$

$$= \frac{1}{\lambda}\sum_{i=1}^{r} S_i - r + \sum_{i=r+1}^{n}\frac{\sum_{x=I_i}^{\infty}\frac{\lambda^{x-1}e^{-\lambda}}{(x-1)!}}{\sum_{x=I_i}^{\infty}\frac{\lambda^{x}e^{-\lambda}}{x\,!}} - (n-r) = 0 \qquad [1]$$

Si definimos

$$P(-1;\lambda) = 0;\, P(A;\lambda) = \sum_{x=0}^{A}\frac{\lambda^{x}e^{-\lambda}}{x\,!}\ (A \geq 0)$$

podemos rescribir [1] en la forma

$$\sum_{i=1}^{r} S_i - \lambda n + \lambda\sum_{i=r+1}^{n}\frac{1-P(I_i-2;\lambda)}{1-P(I_i-1;\lambda)} = 0 \qquad [2]$$

Para obtener λ debe ser resuelta numéricamente la ecuación [2].

Como

$$\frac{1-P(I_i-2;\lambda)}{1-P(I_i-1;\lambda)} > 1\,,$$

entonces

$$\sum_{i=r+1}^{n}\frac{1-P(I_i-2;\lambda)}{1-P(I_i-1;\lambda)} > n-r\,,$$

por lo que

$$\sum_{i=1}^{r} S_i - \lambda n + \lambda(n-r) < 0$$

y, por tanto,

$$\sum_{i=1}^{r} S_i < \lambda r \Rightarrow \lambda > \frac{\sum_{i=1}^{r} S_i}{r}$$

En particular, si $r = n$, o sea si no hay datos censurados, el estimador MV de λ es

$$\sum_{i=1}^{n} S_i - \lambda n = 0; \hat{\lambda} = n^{-1} \sum_{i=1}^{n} S_i = \overline{x}$$

Ejemplo 1: A partir de los datos de la tabla estimar λ.

Ventas	3	9	7	7	8	13	11
Stock disponible	15	12	12	13	13	13	11

Sustituyendo estos valores en la ecuación [2] se tiene

$$34 - \lambda.7 + \lambda \left(\frac{1 - P(11;\lambda)}{1 - P(12;\lambda)} + \frac{1 - P(9;\lambda)}{1 - P(10;\lambda)} \right) = 0$$

Resolviendo la ecuación anterior con *Mathematica* se obtiene $\hat{\lambda} = 8,673$.

3. MÉTODO APROXIMADO DE ESTIMACIÓN

Como alternativa al método de máxima verosimilitud se presenta un método aproximado (MA) de estimación del parámetro λ, basado en la estimación de la demanda insatisfecha esperada. La idea central es similar a la que subyace en los métodos de Bell (1981) y Wecker (1978). Estos autores proponen sendos procedimientos en los que, a partir de la demanda esperada condicionada por el hecho de que ésta es mayor o igual que el stock disponible (para los períodos con datos censurados), se van actualizando los parámetros de la distribución de demanda. La diferencia con el método presentado aquí es que éste sirve para estimar el parámetro y en los mencionados más arriba se debe conocer a priori, precisamente, el valor de ese parámetro.

Sea N_T la variable aleatoria que representa la demanda insatisfecha en un período T (T=r+1,...,n) en el que $S_T = I_T$; es decir, un período en el que no se puede satisfacer la demanda de algunos clientes por haberse agotado el stock disponible. La función de probabilidad de N_T será

$$P(N_T = k) = P(D_T = I_T + k \mid D_T \geq I_T) =$$
$$= \frac{e^{-\lambda} \dfrac{\lambda^{I_T+k}}{(I_T+k)!}}{\sum_{x=I_T}^{\infty} \dfrac{e^{-\lambda}\lambda^x}{x!}} = \frac{1}{1 - P(I_T - 1;\lambda)} e^{-\lambda} \frac{\lambda^{I_T+k}}{(I_T+k)!}; k = 0,1,\ldots$$

La demanda insatisfecha esperada será

$$E[N_T] = \frac{1}{1-P(I_T-1;\lambda)} \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^{I_T+k}}{(I_T+k)!}$$

Haciendo $I_T+k=u$, se tiene

$$\begin{aligned} E[N_T] &= \frac{1}{1-P(I_T-1;\lambda)} \sum_{u=I_T}^{\infty} (u-I_T) e^{-\lambda} \frac{\lambda^u}{u!} = \\ &= \frac{1}{1-P(I_T-1;\lambda)} \left[\sum_{u=I_T}^{\infty} e^{-\lambda} \frac{\lambda^u}{(u-1)!} - I_T \sum_{u=I_T}^{\infty} e^{-\lambda} \frac{\lambda^u}{u!} \right] = \\ &= \frac{\lambda\left[1-P(I_T-2;\lambda)\right] - I_T\left[1-P(I_T-1;\lambda)\right]}{1-P(I_T-1;\lambda)} = \\ &= \frac{\lambda\left[1-P(I_T-1;\lambda)+P(I_T-1;\lambda)-P(I_T-2;\lambda)\right] - I_T\left[1-P(I_T-1;\lambda)\right]}{1-P(I_T-1;\lambda)} \end{aligned}$$

y como

$$\lambda\left[P(I_T-1;\lambda)-P(I_T-2;\lambda)\right] = I_T\left[P(I_T;\lambda)-P(I_T-1;\lambda)\right]$$

entonces

$$\begin{aligned} E[N_T] &= \lambda + \frac{I_T\left[P(I_T;\lambda)-P(I_T-1;\lambda)\right] - I_T\left[1-P(I_T-1;\lambda)\right]}{1-P(I_T-1;\lambda)} = \\ &= \lambda - I_T \frac{1-P(I_T;\lambda)}{1-P(I_T-1;\lambda)} \end{aligned} \qquad [3]$$

Partiendo de la expresión anterior el método aproximado (MA) procede como sigue:

- Paso 1: Se calcula $\lambda_1 = \frac{\sum_{i=1}^{r} S_i}{r}$ y en [3] se sustituye λ por λ_1. Se obtiene así una primera aproximación de la demanda insatisfecha esperada, $\overline{N}_{T1}$, para $T=r+1,...,n$.

- Paso 2: Se recalculan las ventas para los períodos con stockout: $S_{T2} = S_T + \overline{N}_{T1}$, $T=r+1,...n$.

- Paso 3: Se obtiene $\lambda_2 = \frac{\sum_{i=1}^{r} S_i + \sum_{T=r+1}^{n} S_{T2}}{n}$.

- Paso 4: Se repite la segunda parte del paso 1 sustituyendo en la expresión [3] λ por λ_2. De este modo se obtiene una segunda aproximación de la demanda insatisfecha esperada, $\overline{N}_{T2}$, para $T=r+1,...,n$.

- Paso 5: Se repite el paso 2: $S_{T3} = S_T + \overline{N}_{T2}$, $T=r+1,...n$.

- Paso 6: Se obtiene finalmente una estimación aproximada del parámetro λ repitiendo el paso 3:

$$\lambda_3 = \frac{\sum_{i=1}^{r} S_i + \sum_{T=r+1}^{n} S_{T3}}{n}$$

<u>Ejemplo 2</u>: Se aplica a continuación el método aproximado (MA) a los datos del ejemplo 1.

- Paso 1: λ_1=34/5=6,8;

$\overline{N}_{61} = 6,8 - 13\dfrac{1-P(13;6,8)}{1-P(12;6,8)} = 0,794$; análogamente, $\overline{N}_{71} = 0,995$.

- Paso 2: $S_{62} = 13 + 0,794 = 13,794$; $S_{72} = 11 + 0,995 = 11,995$.
- Paso 3: $\lambda_2 = (34 + 13,794 + 11,995)/7 = 8,54$.
- Paso 4: $\overline{N}_{62} = 8,54 - 13\dfrac{1-P(13;8,54)}{1-P(12;8,54)} = 1,158$; $\overline{N}_{72} = 1,478$.
- Paso 5: $S_{63} = 13 + 1,158 = 14,158$; $S_{73} = 11 + 1,478 = 12,478$.
- Paso 6: $\lambda_3 = (34 + 14,158 + 12,478)/7 = 8,662$, valor que se aproxima bastante al obtenido por el método de máxima verosimilitud: 8,673.

4. COMPARACIÓN ENTRE LOS MÉTODOS MV Y MA

Con objeto de comparar los métodos MV y MA se han utilizado datos de ventas de una revista mensual dedicada al mundo del cine para la que las ventas pueden considerarse estacionarias. Se dispone de datos reales de unos 700 puntos de venta de la provincia de Bizkaia, de los que se ha tomado al azar una muestra de tamaño 137. Después de eliminar 37 puntos en los que no había datos suficientes (quioscos que cierran en verano, otros que dejan de vender la publicación, etc.) se obtienen los resultados de la tabla 1. Las variables *mv* y *ma* son las estimaciones de λ proporcionadas por el método de máxima verosimilitud y por el método aproximado, respectivamente. La variable *dif* es $(ma - mv)/mv$ en tanto por ciento, y *stockout* es la proporción de datos censurados (o tasa de períodos fuera de stock). Como se ve en la tabla 2, en el 90% de los casos la tasa de stockouts es menor del 30% (stockouts infrecuentes). Con objeto de analizar la influencia del número de datos disponibles sobre las estimaciones, en unos puntos se han usado datos de 12 períodos y en otros de 24. En todos los casos se ha constatado el ajuste del lado izquierdo de la distribución empírica a una distribución teórica de Poisson.

punto	mv	ma	dif	per	stockout	punto	mv	ma	dif	per	stockout
1	1,054	1,048	-0,61%	12	25,00%	*51*	2,470	2,469	-0,04%	24	12,50%
2	10,265	10,265	0,00%	12	8,33%	*52*	3,493	3,489	-0,11%	24	20,83%
3	3,901	3,895	-0,15%	12	25,00%	*53*	3,912	3,910	-0,05%	24	16,67%
4	18,196	18,182	-0,08%	12	25,00%	*54*	9,757	9,757	0,00%	24	8,33%
5	0,828	0,827	-0,18%	12	16,67%	*55*	10,520	10,516	-0,04%	24	12,50%
6	9,089	9,087	-0,02%	12	16,67%	*56*	11,229	11,228	-0,01%	24	8,33%
7	29,521	29,517	-0,01%	12	16,67%	*57*	2,953	2,951	-0,07%	24	16,67%
8	2,420	2,418	-0,08%	12	25,00%	*58*	25,171	25,171	0,00%	24	4,17%
9	18,513	18,510	-0,02%	12	16,67%	*59*	3,944	3,930	-0,35%	24	29,17%
10	18,618	18,615	-0,02%	12	16,67%	*60*	2,233	2,232	-0,04%	24	12,50%
11	3,891	3,889	-0,05%	12	16,67%	*61*	3,767	3,767	0,00%	24	12,50%
12	16,382	16,382	0,00%	12	8,33%	*62*	9,284	9,283	-0,01%	24	12,50%
13	3,661	3,642	-0,52%	12	33,33%	*63*	10,175	10,175	0,00%	24	8,33%
14	4,801	4,794	-0,15%	12	25,00%	*64*	3,327	3,327	0,00%	24	4,17%
15	2,667	2,667	0,00%	12	0,00%	*65*	2,390	2,390	0,00%	24	12,50%
16	1,781	1,779	-0,11%	12	16,67%	*66*	1,617	1,616	-0,06%	24	12,50%
17	1,083	1,083	0,00%	12	0,00%	*67*	8,320	8,319	-0,01%	24	12,50%
18	7,261	7,243	-0,25%	12	33,33%	*68*	2,354	2,351	-0,13%	24	20,83%
19	22,110	22,107	-0,01%	12	16,67%	*69*	1,472	1,470	-0,14%	24	16,67%
20	1,952	1,939	-0,67%	12	25,00%	*70*	7,381	7,381	0,00%	24	8,33%
21	11,294	11,292	-0,02%	12	16,67%	*71*	1,898	1,879	-1,00%	24	33,33%
22	4,283	4,267	-0,37%	12	33,33%	*72*	0,612	0,611	-0,16%	24	12,50%
23	7,685	7,685	0,00%	12	8,33%	*73*	1,105	1,100	-0,45%	24	25,00%
24	2,121	2,120	-0,05%	12	16,67%	*74*	16,782	16,777	-0,03%	24	20,83%
25	11,437	11,433	-0,03%	12	16,67%	*75*	1,177	1,161	-1,37%	24	33,33%
26	18,795	18,793	-0,01%	12	16,67%	*76*	0,591	0,590	-0,22%	24	16,67%
27	2,374	2,356	-0,76%	12	33,33%	*77*	7,937	7,937	0,00%	24	4,17%
28	8,610	8,608	-0,02%	12	16,67%	*78*	0,412	0,410	-0,40%	24	16,67%
29	1,071	1,069	-0,19%	12	16,67%	*79*	3,447	3,447	0,00%	24	4,17%
30	18,029	18,025	-0,02%	12	16,67%	*80*	1,111	1,105	-0,54%	24	25,00%
31	1,665	1,659	-0,36%	12	25,00%	*81*	20,922	20,919	-0,01%	24	16,67%
32	1,540	1,540	0,00%	12	8,33%	*82*	3,010	3,010	0,00%	24	12,50%
33	5,723	5,723	0,00%	12	8,33%	*83*	9,122	9,116	-0,07%	24	20,83%
34	7,443	7,440	-0,04%	12	16,67%	*84*	0,737	0,737	-0,02%	24	8,33%
35	2,583	2,583	0,00%	12	0,00%	*85*	7,980	7,980	0,00%	24	8,33%
36	0,703	0,688	-2,13%	12	33,33%	*86*	1,116	1,114	-0,17%	24	16,67%
37	13,906	13,904	-0,01%	12	16,67%	*87*	2,533	2,531	-0,08%	24	16,67%
38	2,625	2,625	0,00%	12	8,33%	*88*	4,768	4,760	-0,17%	24	25,00%
39	2,990	2,914	-2,54%	12	50,00%	*89*	1,016	1,005	-1,08%	24	29,17%
40	6,050	6,041	-0,15%	12	25,00%	*90*	10,089	10,089	0,00%	24	8,33%
41	4,917	4,917	0,00%	12	0,00%	*91*	12,694	12,693	-0,01%	24	12,50%
42	4,164	4,164	0,00%	12	8,33%	*92*	9,588	9,588	0,00%	24	4,17%
43	0,666	0,660	-0,90%	12	25,00%	*93*	4,913	4,909	-0,08%	24	20,83%
44	2,195	2,158	-1,69%	12	41,67%	*94*	1,635	1,634	-0,06%	24	8,33%
45	1,942	1,940	-0,10%	12	16,67%	*95*	2,670	2,666	-0,15%	24	20,83%
46	3,250	3,250	0,00%	12	0,00%	*96*	2,213	2,209	-0,18%	24	20,83%
47	1,857	1,823	-1,83%	12	41,67%	*97*	1,401	1,397	-0,29%	24	20,83%
48	2,251	2,245	-0,27%	12	25,00%	*98*	3,152	3,150	-0,06%	24	16,67%
49	0,998	0,996	-0,20%	12	16,67%	*99*	3,000	2,998	-0,07%	24	12,50%
50	2,905	2,905	0,00%	12	8,33%	*100*	1,646	1,635	-0,67%	24	29,17%

Tabla 1

stockout en %	frecuencia	frecuencia acumulada
[0,15]	38	38
(15,30]	52	90
(30,45]	9	99
(45,60]	1	100
(60,75]	0	100
(75,100]	0	100
	100	

Tabla 2

En la tabla 3 y en la figura 1 se observa la distribución de la variable *dif*. Como se puede advertir, en más del 90% de los casos el valor absoluto de esta variable es menor del 0,8%.

dif en %	frecuencia	frecuencia acumulada
[-3,-1]	7	7
(-1,-0.8]	1	8
(-0.8,-0.6]	3	11
(-0.6,-0.4]	5	16
(-0.4,-0.2]	7	23
(-0.2,0]	77	100
	100	

Tabla 3

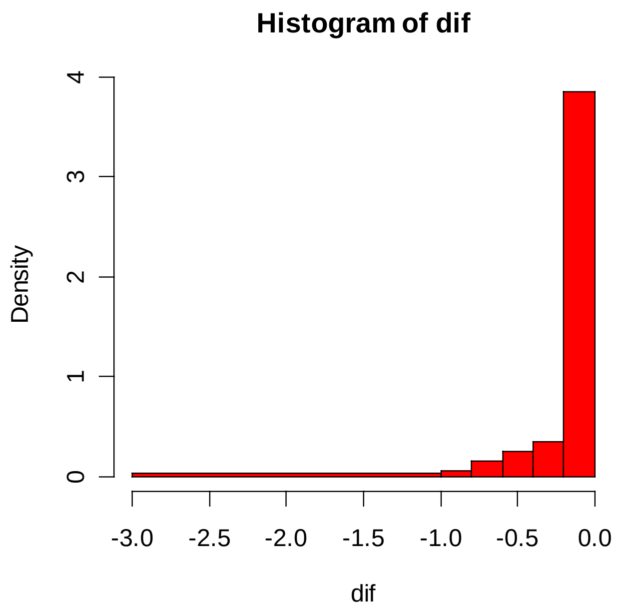

Figura 1

Con objeto de contrastar si existen diferencias significativas entre los valores de la variable *dif* para 12 y 24 períodos, se ha efectuado un test de Wilcoxon, cuyo resultado no permite concluir la existencia de tales diferencias, algo que queda también patente en los diagramas box-plot con muescas de la figura 2.

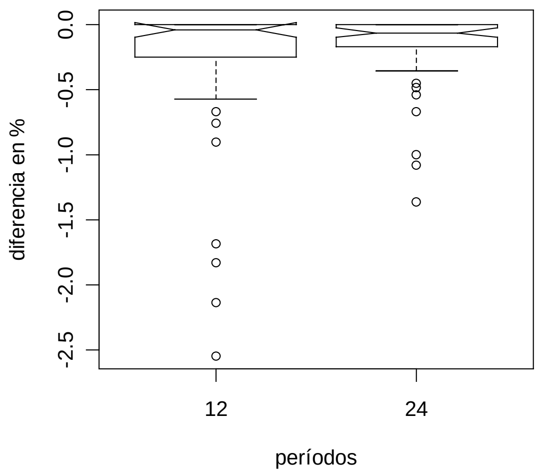

Figura 2

En la figura 3 se representa el diagrama de dispersión entre las variables *dif* y *stockout*. El gráfico sugiere una cierta correlación entre ambas variables**,** de modo que cuanto mayor es el nivel de stockout mayor es la discrepancia entre el método aproximado y la estimación MV, si bien tal discrepancia es inferior al 1% casi en la práctica totalidad de los casos en los que el nivel de censura es menor del 30%.

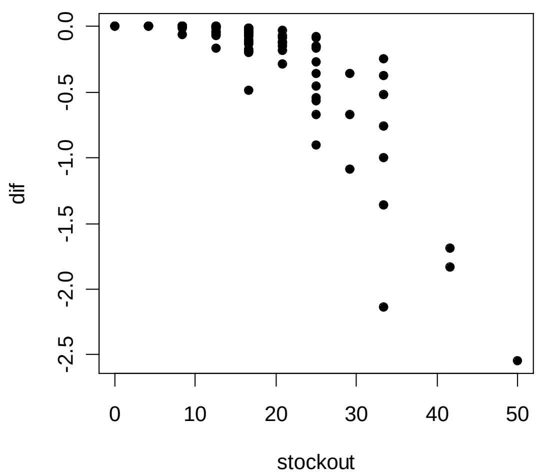

Figura 3

5. CONCLUSIONES

Como se deduce de la prueba presentada en el apartado anterior, ambos métodos, MV y MA, producen resultados prácticamente idénticos. La diferencia máxima entre ambos no excede en ningún caso el 2,54% en valor absoluto, y en el 90% de los casos es menor del 0,8%. Además, parece no tener influencia significativa en esa diferencia la cantidad de datos empleada en la estimación, ni tampoco el nivel de stockout, siempre que éste sea relativamente bajo (menos del 30%).

En cuanto a su puesta en práctica, ambos métodos difieren notablemente. En el método MV se necesita resolver numéricamente una ecuación para cada uno de los puntos, lo cual puede significar un alto requerimiento de software y de tiempo de cómputo. El método MA, por el contrario, es mucho más sencillo de implementar, pues basta una simple hoja de cálculo. Este procedimiento, o una variante suya, sería todavía más ventajoso en el caso de publicaciones diarias, por la inmediatez exigida en la estimación de la demanda.

BIBLIOGRAFÍA

BELL P. (1981): "Adaptative sales forecasting with many stockouts". *Journal of the Operational Research Society* 32, 865-873.

CONRAD S.A. (1976): "Sales data and the estimation of demand". *Operational Research Quarterly* 27, 123-127.

LAU H.-S y LAU A.H.-L. (1996): "Estimating the demand distributions of single-period items having frequent stockouts". *European Journal of Operational Research*, 92, 254-265.

WECKER W. (1978): "Predicting demand from sales data in the presence of stockouts". *Management Science* 24, 1043-1054.