

Aditiva: Variación estacional constante  
 Multiplicativa: Variación estacional creciente/decreciente

## Métodos de descomposición

La aplicación de los **modelos de descomposición** para pronosticar series temporales que manifiestan **tendencia y efectos estacionales** se estudia en este capítulo. No hay bases teóricas para dichos modelos —son rigurosamente un enfoque intuitivo. No obstante, estos modelos han resultado **útiles** cuando los **parámetros que describen una serie temporal no cambian en el tiempo**. La idea básica en la que se apoyan los modelos es la descomposición de las series temporales en varios **factores: tendencia, (error) estacional, cíclico e irregular**. Las estimaciones de estos factores se utilizan para describir las series temporales. Además, si los parámetros de las series temporales no cambian, las estimaciones se pueden usar para determinar pronósticos puntuales.

En la Sección 7.1 se trata el **modelo de descomposición multiplicativa**. Este modelo es útil al modelar series temporales que manifiestan una **variación estacional creciente o decreciente**. En la Sección 7.2 se explica brevemente el **modelo de descomposición aditiva**, que se puede utilizar para modelar series temporales que muestran una **variación estacional constante**. Una breve explicación de X-12-ARIMA, un método de descomposición que aplica la Oficina de Censos del Departamento de Comercio de Estados Unidos se presenta en la Sección 7.3.

factores ST:

- + Tendencia
- + estacional
- + cíclico
- + irregular

se estiman  
para describir  
ST

parámetros para  
describir una serie  
temporal:

Si los  
no cambian  
las estimaciones  
de series  
fueron buenas

## 7.1 DESCOMPOSICIÓN MULTIPLICATIVA

Considere una serie temporal que manifiesta **variación estacional creciente** o decreciente. Cuando los parámetros que describen la serie no cambian en el tiempo, a veces, la serie temporal se puede modelar en forma adecuada usando lo que se llama **modelo de descomposición multiplicativa**. Este modelo se puede plantear como sigue:

El modelo de descomposición multiplicativa es

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

donde

$y_t$  = valor observado de la serie temporal en el periodo  $t$

$TR_t$  = el **componente (o factor) de la tendencia** en el periodo  $t$

$SN_t$  = el **componente (o factor) estacional** en el periodo  $t$

$CL_t$  = el **componente (o factor) cíclico** en el periodo  $t$

$IR_t$  = el **componente (o factor) irregular** en el periodo  $t$

Ya se trató la naturaleza de los efectos de la tendencia, de las variaciones estacionales y de las fluctuaciones irregulares. El componente cíclico,  $CL_t$ , se refiere a los movimientos recurrentes hacia arriba y hacia abajo con respecto a los niveles de la tendencia, cuando los causa, por ejemplo, el ciclo mercantil. Estas fluctuaciones pueden durar desde dos a más de diez años, según lo que se mide de máximo a máximo o de mínimo a mínimo. En las cuestiones de los negocios, un máximo marcaría el final de una expansión en la actividad comercial, y un mínimo señalaría el final de una contracción.

Obsérvese que en este modelo de descomposición se utiliza un **factor estacional multiplicativo**. Es decir, el factor estacional se multiplica por la tendencia (y no se suma a la tendencia como en la regresión con variable ficticia; véase Sección 6.4). Para ver la manera como el factor estacional multiplicativo puede modelar la **variación estacional creciente**, suponga que, por decir algo, las ventas de los motores fuera de borda que ofrece Power Drive Corporation son estacionales. También suponga que las ventas más bajas son las del primer trimestre, las más altas se observan en el segundo trimestre, las ventas moderadamente altas se presentan en el tercer trimestre y las moderadamente bajas en el cuarto trimestre. Además, suponga que las ventas manifiestan una **tendencia lineal** dada por

$$TR_t = 500 + 50t$$

donde  $t = 10$  se considera como el cuarto trimestre de 2002. Si se considera sólo la tendencia, es de esperarse que las ventas de los motores fuera de borda en los cuatro trimestres de 2003 sean

$$TR_1 = 500 + 50(1) = 550 \quad (\text{trimestre 1})$$

$$TR_2 = 500 + 50(2) = 600 \quad (\text{trimestre 2})$$

$$TR_3 = 500 + 50(3) = 650 \quad (\text{trimestre 3})$$

$$TR_4 = 500 + 50(4) = 700 \quad (\text{trimestre 4})$$

Sin embargo, las ventas son estacionales. Por lo tanto, podemos modelar el comportamiento estacional de las ventas mediante la definición de factores estacionales. Suponga ahora que los factores estacionales para los trimestres 1, 2, 3 y 4 son  $SN_{Q1} = .4$ ,  $SN_{Q2} = 1.6$ ,  $SN_{Q3} = 1.2$ , y  $SN_{Q4} = .8$ . Si suponemos que estos factores estacionales son multiplicativos, entonces cuando consideramos tanto la tendencia como los efectos estacionales, esperamos que las ventas en los cuatro trimestres de 2003 serán de

$$TR_1 \times SN_{Q1} = [500 + 50(1)](.4) = 220$$

$$TR_2 \times SN_{Q2} = [500 + 50(2)](1.6) = 960$$

$$TR_3 \times SN_{Q3} = [500 + 50(3)](1.2) = 780$$

$$TR_4 \times SN_{Q4} = [500 + 50(4)](.8) = 560$$

Al multiplicar la tendencia  $TR_t$  por los factores estacionales apropiados se modela el patrón estacional de ventas. Esto se ilustra en la Figura 7.1.

Si los factores estacionales multiplicativos son constantes en el tiempo, entonces podemos modelar la variación estacional creciente. Por ejemplo, en la situación de las ventas de los motores fuera de borda, cuando tomamos en cuenta tanto la tendencia como los efectos estacionales, es de esperar que las ventas en los cuatro trimestres de 2004 sean

$$TR_5 \times SN_{Q1} = [500 + 50(5)](.4) = 300$$

$$TR_6 \times SN_{Q2} = [500 + 50(6)](1.6) = 1280$$

$$TR_7 \times SN_{Q3} = [500 + 50(7)](1.2) = 1020$$

$$TR_8 \times SN_{Q4} = [500 + 50(8)](.8) = 720$$

La multiplicación de la tendencia por los factores estacionales significa que las dimensiones de la variación estacional serán proporcionales a la tendencia. Por lo tanto, puesto que la tendencia

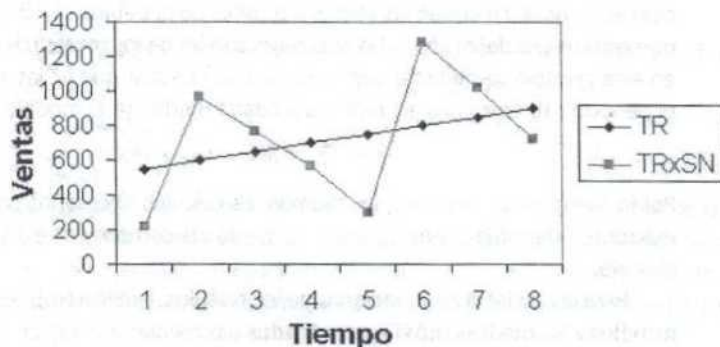
$$TR_t = 500 + 50t$$

es creciente, las dimensiones de la variación estacional son crecientes (variación estacional creciente). Tome en cuenta otra vez que estamos suponiendo que los factores estacionales son constantes en el tiempo. A veces, los factores estacionales sí cambian en el tiempo; entonces se tienen que recurrir a los métodos de los Capítulos 8 al 12.

El factor estacional  $SN_t$  modela los patrones cíclicos en una serie temporal que son completados dentro de un año civil. Si una serie temporal manifiesta un ciclo

**FIGURA 7.1**

Ejemplo obtenido con Excel de los factores estacionales multiplicativos  $SN_{Q1} = .4$ ,  $SN_{Q2} = 1.6$ ,  $SN_{Q3} = 1.2$ , y  $SN_{Q4} = .8$





que tiene una mayor duración, se puede definir un factor cíclico  $CL_t$ . Por ejemplo, en las ventas de los motores fuera de borda, suponga que los cuatro trimestres de 2003 están incluidos en un “periodo de auge” del ciclo comercial. Suponga que los factores cíclicos que describen la actividad económica incrementada en los cuatro trimestres de 2003 son  $CL_1 = 1.08$ ,  $CL_2 = 1.09$ ,  $CL_3 = 1.09$ , y  $CL_4 = 1.10$ . Si se toman en cuenta los efectos de la tendencia, de la variación estacional y los cíclicos, es de esperar que las ventas en los cuatro trimestres de 2003 sean

$$TR_1 \times SN_{Q1} \times CL_1 = 220(1.08) = 238$$

$$TR_2 \times SN_{Q2} \times CL_2 = 960(1.09) = 1046$$

$$TR_3 \times SN_{Q3} \times CL_3 = 780(1.09) = 850$$

$$TR_4 \times SN_{Q4} \times CL_4 = 560(1.10) = 616$$

Por consiguiente, los factores cíclicos incrementan las ventas esperadas por arriba de niveles que se podrían esperar si sólo se consideraran los efectos de la tendencia y los estacionales. Esto refleja la prosperidad repentina en la actividad económica.

El **método de descomposición multiplicativa** se puede aplicar para obtener estimaciones puntuales —denotadas por  $tr_t$ ,  $sn_t$ ,  $cl_t$  e  $ir_t$ — de los factores  $TR_t$ ,  $SN_t$ ,  $CL_t$  e  $IR_t$ . El ejemplo siguiente ilustra el procedimiento.

### EJEMPLO 7.1

Discount Soda Shop, Inc. es propietaria y administra diez establecimientos donde se venden bebidas refrescantes a los automovilistas sin que éstos tengan que descender de sus vehículos. Discount Soda vende Tasty Cola, una bebida con agua carbonatada que fue introducida al mercado hace tres años y que ha ganado aceptación. Cada cierto tiempo, Discount Soda hace un pedido de Tasty Cola al distribuidor regional. La compañía utiliza una estrategia para su inventario con la que pretende cumplir prácticamente con toda la demanda de Tasty Cola, a la vez que asegura que la compañía no inmoviliza sin necesidad su dinero, pidiendo mucho más Tasty Cola de la que espera razonablemente vender. Con objeto de poner en marcha su política de inventario, Discount Soda necesita tener pronósticos mensuales de las ventas de Tasty Cola (en cientos de envases). Al final de cada mes, Discount Soda desea pronósticos puntuales y pronósticos de los intervalos de predicción de las ventas de Tasty Cola en los meses futuros.

Discount Soda lleva los registros de las ventas mensuales de las ventas de Tasty Cola durante los tres años anteriores, a los que se les puede llamar año 1, año 2 y año 3. Esta serie temporal se ilustra en la Tabla 7.1 y se grafica en la Figura 7.2. Observe que aparte de mostrar una tendencia lineal, la serie temporal de las ventas de Tasty Cola posee **variación estacional**, en la que las ventas más altas de la bebida se observan en el verano y los primeros meses del otoño, y las más bajas son las de los meses del invierno. Más adelante en este ejemplo se muestra que es razonable concluir que  $y_t$ , las ventas de Tasty Cola en el periodo  $t$  se describen en forma adecuada mediante el modelo

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

Por lo tanto, en la Tabla 7.2 se resumen los cálculos necesarios para determinar las estimaciones —denotadas por  $tr_t$ ,  $sn_t$ ,  $cl_t$  e  $ir_t$ — de los componentes  $TR_t$ ,  $SN_t$ ,  $CL_t$  e  $IR_t$  de este modelo.

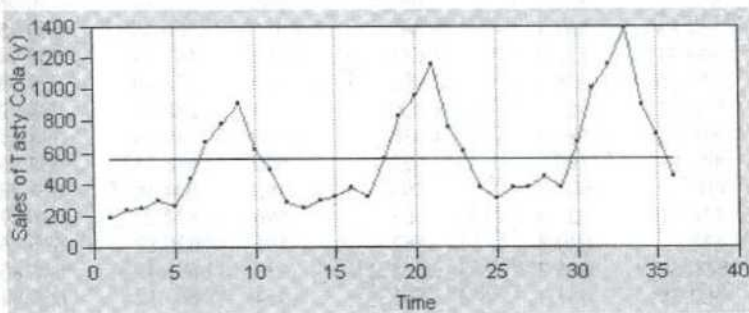
Para iniciar las consideraciones de los cálculos, explicaremos el cálculo de las **medias móviles** y las **medias móviles centradas** (las medias móviles centradas, se denotan con

TABLA 7.1 Ventas mensuales de Tasty Cola (en cientos de envases)

Ventas				Ventas			
Año	Mes	t	$y_t$	Año	Mes	t	$y_t$
1	1 (Ene.)	1	189	2	7	19	831
	2 (Feb.)	2	229		8	20	960
	3 (Mar.)	3	249		9	21	1152
	4 (Abr.)	4	289		10	22	759
	5 (May.)	5	260		11	23	607
	6 (Jun.)	6	431		12	24	371
	7 (Jul.)	7	660	3	1	25	298
	8 (Ags.)	8	777		2	26	378
	9 (Sept.)	9	915		3	27	373
	10 (Oct.)	10	613		4	28	443
	11 (Nov.)	11	485		5	29	374
	12 (Dic.)	12	277		6	30	660
2	1	13	244		7	31	1004
	2	14	296		8	32	1153
	3	15	319		9	33	1388
	4	16	370		10	34	904
	5	17	313		11	35	715
	6	18	556		12	36	441

FIGURA 7.2

Gráfica JMP IN de las ventas mensuales de Tasty Cola (en cientos de envases)



CMA<sub>1</sub>). El objetivo que se oculta detrás del cálculo de estas medias es eliminar las variaciones estacionales y las fluctuaciones irregulares de los datos. La primera media móvil es el promedio de los primeros 12 valores de las ventas de Tasty Cola

$$\frac{189 + 229 + 249 + 289 + 260 + 431 + 660 + 777 + 915 + 613 + 485 + 277}{12} = 447.833$$

Aquí usamos una "media móvil de 12 periodos" porque los datos de la serie temporal de Tasty Cola son mensuales ( $L = 12$  periodos o "estaciones" por año). Si los datos fueran de trimestres, calcularíamos una "media móvil de cuatro periodos". La segunda media móvil se obtiene al eliminar el primer valor de las ventas ( $y_1$ ) del promedio e incluir el siguiente



**TABLA 7.2** Análisis de la serie temporal de las ventas anteriores de Tasty Cola, por medio de la descomposición multiplicativa

(a) Valores de  $sn_t$ ,  $d_t$  y  $tr_t$

	A	B	C	D	E	F	G	H
	Media móvil							
t	$y_t$	de 12 Periodos	$CMA_t = tr_t \times cl_t$	$sn_t \times ir_t = y_t / (tr_t \times cl_t)$	$sn_t$	$d_t = \frac{y_t}{sn_t}$	$tr_t = 380.163 + 9.489t$	$\hat{y}_t = tr_t \times sn_t$
1	189				.493	383.37	389.652	192.10
2	229				.596	384.23	399.141	237.89
3	249				.595	418.49	408.630	243.13
4	289				.680	425	418.119	284.32
5	260				.564	460.99	427.608	241.17
6	431				.986	437.12	437.097	430.98
7	660	447.833	450.1	1.466	1.467	449.9	446.586	655.14
8	777	452.417	455.2	1.707	1.693	458.95	456.075	772.13
9	915	458	460.9	1.985	1.990	459.79	465.564	926.47
10	613	463.833	467.2	1.312	1.307	469.01	475.053	620.89
11	485	470.583	472.8	1.026	1.029	471.33	489.542	498.59
12	277	475	480.2	.577	.600	461.67	494.031	296.42
13	244	485.417	492.5	.495	.493	494.97	503.520	248.24
14	296	499.667	507.3	.583	.596	496.64	513.009	305.75
15	319	514.917	524.8	.608	.595	536.13	522.498	310.89
16	370	534.667	540.7	.684	.680	544.12	531.987	361.75
17	313	546.833	551.9	.567	.564	554.97	541.476	305.39
18	556	557	560.9	.991	.986	563.89	550.965	543.25
19	831	564.833	567.1	1.465	1.467	566.46	560.454	822.19
20	960	569.333	572.7	1.676	1.693	567.04	569.943	964.91
21	1152	576.167	578.4	1.992	1.990	578.89	579.432	1153.07
22	759	580.667	583.7	1.300	1.307	580.72	588.921	769.72
23	607	586.75	589.3	1.030	1.029	589.89	598.410	615.76
24	371	591.833	596.2	.622	.600	618.33	607.899	364.74
25	298	600.5	607.7	.490	.493	604.46	617.388	304.37
26	378	614.917	623.0	.607	.596	634.23	626.877	373.62
27	373	631	640.8	.582	.595	626.89	636.366	378.64
28	443	650.667	656.7	.675	.680	651.47	645.855	439.18
29	374	662.75	667.3	.561	.564	663.12	655.344	369.61
30	660	671.75	674.7	.978	.986	669.37	664.833	655.53
31	1004	677.583			1.467	684.39	674.322	989.23
32	1153				1.693	681.04	683.811	1157.69
33	1388				1.990	697.49	693.300	1379.67
34	904				1.307	691.66	702.789	918.55
35	715				1.029	694.85	712.278	732.93
36	441				.600	735	721.707	433.06

valor de las ventas ( $y_{13}$ ) en el promedio. Por consiguiente, se obtiene

$$229 + 249 + 289 + 260 + 431 + 660 + 777 + 915 + 613 + 485 + 277 + 244$$

12

$$= 452.417$$

TABLA 7.2 (Continuación)

(b) Valores de  $cl_t$  e  $ir_t$ 

$t$	$y_t$	$tr_t \times sn_t$	$cl_t \times ir_t = \frac{y_t}{tr_t \times sn_t}$	$cl_t = \frac{cl_{t-2} ir_{t-2} 1}{3}$	$ir_t = \frac{cl_t \times ir_t}{cl_t}$
1	189	192.10	.9839		
2	229	237.89	.9626		
3	249	243.13	1.0241		
4	289	284.32	1.0165		
5	260	241.17	1.0781		
6	431	430.98	1.0000		
7	660	655.14	1.0074		
8	777	772.13	1.0063		
9	915	926.47	.9876		
10	613	620.89	.9873		
11	485	498.59	.9727		
12	277	296.42	.9345		
13	244	248.24	.9829		
14	296	305.75	.9681		
15	319	310.89	1.0261		
16	370	361.75	1.0228		
17	313	305.39	1.0249		
18	556	543.25	1.0235		
19	831	822.19	1.0107		
20	960	964.91	.9949		
21	1152	1153.07	.9991		
22	759	769.72	.9861		
23	607	615.76	.9858		
24	371	364.74	1.0172		
25	298	304.37	.9791		
26	378	373.62	1.0117		
27	373	378.64	.9851		
28	443	439.18	1.0087		
29	374	369.61	1.0119		
30	660	655.53	1.0068		
31	1004	989.23	1.0149		
32	1153	1157.69	.9959		
33	1388	1379.67	1.0060		
34	904	918.55	.9842		
35	715	732.93	.9755		
36	441	433.06	1.0183		

La tercera media móvil se obtiene eliminando  $y_2$  del promedio e incorporando  $y_{14}$  al promedio. Tenemos entonces

$$\frac{249 + 289 + 260 + 431 + 660 + 777 + 915 + 613 + 485 + 277 + 244 + 296}{12}$$

$$= 458$$

Las sucesivas medias móviles se calculan en forma similar hasta que no quede incluida  $y_{36}$  en el último promedio móvil. Obsérvese que usamos aquí el término "media móvil" porque al calcular estos promedios, avanzamos eliminando la observación más remota en la media anterior e incluyendo la observación "siguiente" en el promedio nuevo.

La primera media móvil corresponde a un tiempo que está en un punto intermedio entre los periodos 6 y 7, la segunda media móvil corresponde a un tiempo que está en el punto medio de los periodos 7 y 8, y así sucesivamente. Con objeto de obtener promedios que correspondan a periodos en la serie temporal original de Tasty Cola, calculamos las **medias móviles centradas**. Los promedios son las medias móviles de dos periodos de las medias móviles de 12 periodos previamente calculadas. Por consiguiente, la primera media móvil centrada es

$$ECM_5 = \frac{447.833 + 452.417}{2} = 450.1$$

La segunda media móvil centrada es

$$\frac{452.417 + 458}{2} = 455.2$$

Las medias móviles centradas sucesivas se determinan de manera similar. Las medias móviles de 12 periodos y las medias móviles centradas para la serie temporal de las ventas de Tasty Cola se dan en la Tabla 7.2(a)

Si las medias móviles originales se determinan usando un número impar de valores de la serie temporal, el procedimiento de centrado sería necesario. Por ejemplo, si tuviéramos tres estaciones por año, entonces calcularíamos tres medias móviles de tres periodos. Luego la primera media móvil correspondería al periodo 2, la segunda media móvil correspondería al periodo 3 y así sucesivamente. Sin embargo, la mayor parte de las series temporales son trimestrales, mensuales, o semanales, por lo que el procedimiento de centrado es necesario.

Se considera que la media móvil centrada en el periodo  $t$ ,  $CMA_t$ , es igual a  $tr_t \times cl_t$ , la estimación de  $TR_t \times CL_t$ . La razón es que se supone que el procedimiento de obtener las medias elimina 1) las variaciones estacionales (obsérvese que cada media móvil se calcula usando exactamente una observación de cada estación) y 2) las fluctuaciones irregulares de corto plazo. Se conservan los efectos de la tendencia (plazo más largo) y los efectos cíclicos, es decir,  $tr_t \times cl_t$ .

Puesto que el modelo

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

quiere decir que

$$SN_t \times IR_t = \frac{y_t}{TR_t \times CL_t} = \frac{y_t}{CMA_t}$$

se infiere que la estimación  $sn_t \times ir_t$  de  $SN_t \times IR_t$  es

$$sn_t \times ir_t = \frac{y_t}{tr_t \times cl_t} = \frac{y_t}{CMA_t}$$



**TABLA 7.3** Estimaciones de los factores estacionales de la serie temporal de las ventas de Tasty Cola

		$sn_t \times ir_t = y_t / (tr_t \times cl_t)$		$\bar{sn}_t$ - promedio mensual	$sn_t = 1.0008758 (\bar{sn}_t)$
		Año 1	Año 2		
1	Ene.	.495	.490	.4925	.493
2	Feb.	.583	.607	.595	.596
3	Mar.	.608	.582	.595	.595
4	Abr.	.684	.675	.6795	.680
5	May	.567	.561	.564	.564
6	Jun	.991	.978	.9845	.986
7	Jul	1.466	1.465	1.4655	1.467
8	Ags.	1.707	1.676	1.6915	1.693
9	Sept.	1.985	1.992	1.9885	1.990
10	Oct.	1.312	1.300	1.306	1.307
11	Nov.	1.026	1.030	1.028	1.029
12	Dic.	.577	.622	.5995	.600

Los valores de  $sn_t \times ir_t$  están ya calculados en la Tabla 7.2(a), por lo que ya podemos encontrar  $sn_t$  agrupando los valores de  $sn_t \times ir_t$ , por meses y calcular el promedio,  $\bar{sn}_t$  de cada mes. Estos factores estacionales se normalizan entonces de modo que se añaden a  $L = 12$ , en número de periodos en el año. Esta normalización se logra multiplicando cada valor de  $\bar{sn}_t$  por la cantidad

$$\frac{L}{\sum_{t=1}^L \bar{sn}_t} = \frac{12}{11.9895} = 1.0008758$$

Este proceso de normalización da como resultado la estimación  $sn_t = 1.0008758 (\bar{sn}_t)$ , que es la estimación de  $SN_t$ . Estos cálculos se resumen en la Tabla 7.3.

Luego de calcular los valores de  $sn_t$  y de colocarlos en la Tabla 7.2(a), definimos la observación compensada respecto a la variación estacional en el periodo  $t$  como

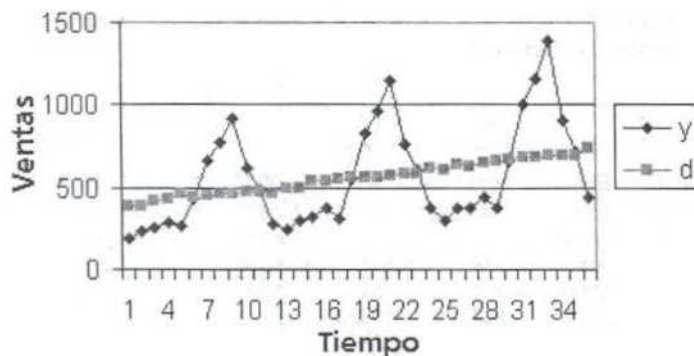
$$d_t = \frac{y_t}{sn_t}$$

El objeto de calcular las observaciones compensadas respecto a la variación estacional es estimar mejor el componente de la tendencia  $TR_t$ . Al dividir  $y_t$  entre el factor estacional estimado se elimina la estacionalidad de los datos y nos permite entender mejor la naturaleza de la tendencia. Las observaciones compensadas respecto a la variación estacional están ya calculadas en la Tabla 7.2(a) y se grafican en la Figura 7.3. Puesto que las observaciones compensadas respecto a la variación estacional muestran una apariencia recta, parece razonable suponer una tendencia lineal

$$TR_t = \beta_0 + \beta_1 t$$

**FIGURA 7.3**

Gráfica de Excel de las ventas compensadas respecto a la variación estacional  $d_t$ .



Estimamos  $TR_t$  ajustando una recta a los datos compensados respecto a la variación estacional. Es decir, calculamos las estimaciones puntuales de mínimos cuadrados de los parámetros que están en el modelo de regresión lineal simple que relaciona la variable dependiente  $d_t$  con la variable independiente  $t$ :

$$d_t = \beta_0 + \beta_1 t + \varepsilon_t$$

Por consiguiente, obtenemos  $tr_t$ , la estimación de  $TR_t$ , al calcular

$$b_1 = \frac{SS_{td}}{SS_{tt}} = \frac{\sum_{t=1}^{36} t d_t - \frac{\left(\sum_{t=1}^{36} t\right) \left(\sum_{t=1}^{36} d_t\right)}{36}}{\sum_{t=1}^{36} t^2 - \frac{\left(\sum_{t=1}^{36} t\right)^2}{36}} = 9.489$$

y

$$b_0 = \bar{d} - b_1 \bar{t} = \frac{\sum_{t=1}^{36} d_t}{36} - b_1 \frac{\sum_{t=1}^{36} t}{36} = 380.163$$

Por lo tanto,

$$tr_t = b_0 + b_1 t = 380.163 + 9.489 t$$

Los valores de  $tr_t$  están ya calculados en la Tabla 7.2(a). Por ejemplo, observe que aunque  $y_{22} = 759$  (ventas de Tasty Cola en el periodo 22) es mayor que  $tr_{22} = 588.921$  (la tendencia estimada en el periodo 22),  $d_{22} = 580.72$  es menor que  $tr_{22} = 588.921$ . Esto significa que según los datos compensados respecto a la variación estacional, las ventas de Tasty Cola estuvieron ligeramente abajo en octubre del año 2. La causa podría haber sido un octubre ligeramente más frío que lo normal.

Por consiguiente, hemos encontrado las estimaciones  $sn_t$  y  $tr_t$  de  $SN_t$  y  $TR_t$ . Como el modelo

$$y_t = TR_t \times SN_t \times CL_t \times IR_t$$

quiere decir que

$$CL_t \times IR_t = \frac{y_t}{TR_t \times SN_t}$$

se infiere que la estimación de  $CL_t \times IR_t$  es

$$cl_t \times ir_t = \frac{y_t}{tr_t \times sn_t}$$

Además, la experiencia demuestra que al considerar datos mensuales o trimestrales podemos promediar  $ir_t$  y, por consiguiente, calcular la estimación  $cl_t$  de  $CL_t$  mediante

$$cl_t = \frac{cl_{t-1}ir_{t-1} + cl_tir_t + cl_{t+1}ir_{t+1}}{3}$$

Es decir,  $cl_t$  es una media móvil de tres periodos de los valores  $cl_t \times ir_t$ .

Por último, determinamos la estimación  $ir_t$  o  $IR_t$  usando la ecuación

$$ir_t = \frac{cl_t \times ir_t}{cl_t}$$

Los cálculos de los valores de  $cl_t$  e  $ir_t$  para los datos de Tasty Cola se resumen en la Tabla 7.2(b). Puesto que sólo hay tres años de datos, y como la mayor parte de los valores de  $cl_t$  está cerca de 1, no podemos percibir un ciclo muy bien definido. Además, al examinar los valores de  $ir_t$ , no podemos detectar un patrón en las estimaciones de los factores irregulares.

Por costumbre, las estimaciones  $tr_t$ ,  $sn_t$ ,  $cl_t$  e  $ir_t$  obtenidas mediante el método de descomposición multiplicativa se utilizan en la descripción de series temporales. No obstante, también se pueden aplicar estas estimaciones para pronosticar valores futuros de las series temporales. Si no hay patrón en el componente irregular, entonces el pronóstico será que  $IR_t$  es igual a 1. Por lo tanto, el pronóstico puntual de  $y_t$  es

$$\hat{y}_t = tr_t \times sn_t \times cl_t$$

si existe un ciclo muy bien definido y puede ser predicho. El pronóstico puntual es

$$\hat{y}_t = tr_t \times sn_t$$

si no existe un ciclo muy bien definido o si  $CL_t$  no puede ser predicho. Por lo que se refiere al ejemplo de Tasty Cola, donde

$$tr_t = b_0 + b_1t = 380.163 + 9.489t$$

los pronósticos puntuales de las  $n = 36$  ventas anteriores de Tasty Cola se proporcionan en la Tabla 7.2(a). Los pronósticos puntuales de las ventas futuras de Tasty Cola en los 12 meses del año 4 se dan en la Tabla 7.4. Por ejemplo, el pronóstico puntual de las ventas en el periodo 44 es

$$\begin{aligned}\hat{y}_{44} &= tr_{44} \times sn_{44} \\ &= [380.163 + 9.489(44)](1.693) = 797.699 (1.693) \\ &= 1350.50\end{aligned}$$

Aunque no hay intervalo de predicción teóricamente correcto de  $y_t$ , los autores encontraron que un intervalo de predicción  $100(1 - \alpha)\%$  claramente preciso (aproximado) para  $y_t$  es

$$[\hat{y}_t \pm B_t[100(1 - \alpha)]]$$

Pronósticos.



**TABLA 7.4** Pronósticos de los valores futuros de las ventas de Tasty Cola calculados usando la descomposición multiplicativa

$-t$	$sn_t$	$tr_t = 380.163 + 9.489t$	$\hat{y}_t = tr_t \times sn_t$	$B_t(95)$	$[\hat{y}_t - B_t(95), \hat{y}_t + B_t(95)]$	$y_t$
37	.493	731.273	360.52	26.80	[333.72, 387.32]	352
38	.596	740.762	441.48	26.92	[414.56, 468.40]	445
39	.595	750.252	446.40	27.04	[419.36, 473.44]	453
40	.680	759.741	516.62	27.17	[489.45, 543.79]	541
41	.564	769.231	433.85	27.30	[406.55, 461.15]	457
42	.986	778.720	767.82	27.44	[740.38, 795.26]	762
43	1.467	788.209	1156.30	27.59	[1128.71, 1183.89]	1194
44	1.693	797.699	1350.50	27.74	[1322.76, 1378.24]	1361
45	1.990	807.188	1606.30	27.89	[1578.41, 1634.19]	1615
46	1.307	816.678	1067.40	28.05	[1039.35, 1095.45]	1059
47	1.029	826.167	850.12	28.22	[821.90, 878.34]	824
48	.600	835.657	501.39	28.39	[473, 529.78]	495

donde  $B_t [100(1 - \alpha)]$  es el límite de error en un intervalo de predicción  $100(1 - \alpha)\%$

$$[tr_t \pm B_t [100(1 - \alpha)]]$$

para la observación compensada respecto a la variación estacional

$$\begin{aligned} d_t &= TR_t + \varepsilon_t \\ &= \beta_0 + \beta_1 t + \varepsilon_t \end{aligned}$$

Por ejemplo, si se usa SAS para predecir  $d_t$  con base en  $t$  usando la línea de tendencia anterior, tenemos que un intervalo de predicción del 95% para  $d_{44}$  es

$$[769.959, 825.439]$$

Esto quiere decir que

$$\begin{aligned} B_{44}[95] &= \frac{825.439 - 769.959}{2} \\ &= 27.74 \end{aligned}$$

Se infiere entonces que un intervalo de predicción aproximado del 95% para  $y_{44}$  es

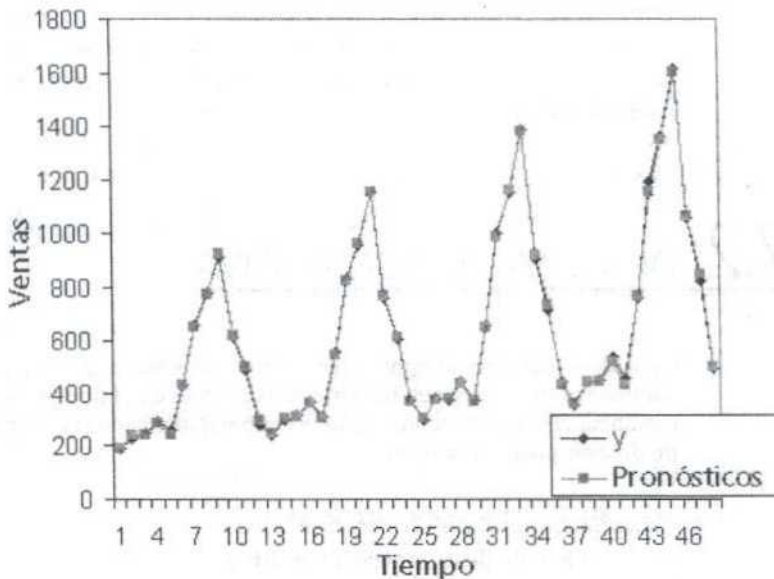
$$[1350.50 - 27.74, 1350.50 + 27.74] = [1322.76, 1378.24]$$

En la Tabla 7.4 se presentan intervalos de predicción del 95% (calculados con el método anterior) para las ventas de Tasty Cola en los 12 meses del año 4.

Suponga que, en realidad, observamos las ventas de Tasty Cola en el año 4 y que estas ventas son las que se indican en la Tabla 7.4. En la Figura 7.4 se grafican las ventas observadas y las pronosticadas de los 48 periodos de venta. En la práctica, el analista usaría la comparación de las ventas observadas y pronosticadas de los años 1 a 3 para determinar si la ecuación de los pronósticos se ajusta en forma conveniente a los datos anteriores. Un ajuste adecuado (como el que se indica en la Figura 7.4, por ejemplo) podría impulsar a un analista a usar esta ecuación para calcular los pronósticos de periodos futuros.

**FIGURA 7.4**

Pronósticos de los valores anteriores y futuros de las ventas de Tasty Cola calculados usando la descomposición multiplicativa



Una razón de que la ecuación del pronóstico de Tasty Cola

$$\begin{aligned}\hat{y}_t &= tr_t \times sn_t \\ &= (380.163 + 9.489t)sn_t\end{aligned}$$

proporcione pronósticos razonables es que esta ecuación *multiplica*  $sn_t$  por  $tr_t$ . Por lo tanto, cuando se incrementa el nivel promedio de la serie temporal (determinado por la tendencia), aumenta el cambio estacional de la serie temporal, lo cual es consecuente con las gráficas de los datos de las Figuras 7.2 y 7.4. Por ejemplo, observe que en la Tabla 7.3 el factor estacional estimado para agosto es 1.693. La ecuación de pronósticos genera una predicción de las ventas de Tasty Cola en agosto del año 1 igual a

$$\begin{aligned}\hat{y}_8 &= [380.163 + 9.489(8)](1.693) \\ &= (456.075)(1.693) \\ &= 772.13\end{aligned}$$

Esto significa un cambio estacional de  $772.13 - 456.075 = 316.055$  (cientos de envases) arriba de 456.075, la tendencia estimada. La ecuación del pronóstico da una predicción para las ventas de Tasty Cola para agosto del año 2 de

$$\begin{aligned}\hat{y}_{20} &= [380.163 + 9.489(20)](1.693) \\ &= (569.943)(1.693) \\ &= 964.91\end{aligned}$$

lo cual lleva a un cambio estacional *incrementado* de  $964.91 - 569.943 = 394.967$  (cientos de envases) por arriba de 569.943, la tendencia estimada. En general, la ecuación del pronóstico es apropiada para pronosticar una serie temporal con un cambio estacional

que es proporcional al nivel promedio de la serie temporal según lo determina la tendencia —es decir, una serie temporal que manifiesta variación estacional creciente. De hecho, la variación estacional creciente recibe el nombre algunas veces de **variación estacional multiplicativa**.

## 7.2 DESCOMPOSICIÓN ADITIVA

Considere una serie temporal que exhibe variación estacional constante. Cuando los parámetros que describen las series no cambian en el tiempo, algunas veces se puede modificar en forma adecuada la serie temporal mediante lo que se conoce como **modelo de descomposición aditiva**.

El modelo de descomposición aditiva es

$$y_t = \underbrace{TR_t}_{\text{Tendencia}} + \underbrace{SN_t}_{\text{factor estacional}} + \underbrace{CL_t}_{\text{factor cíclico}} + \underbrace{IR_t}_{\text{factor irregular}}$$

En este caso,  $TR_t$ ,  $SN_t$ ,  $CL_t$  e  $IR_t$  están definidas como la tendencia, factor estacional, factor cíclico y factor irregular. Sin embargo, estos factores son aditivos en lugar de multiplicativos.

El **método de descomposición aditiva** se usa para determinar estimaciones puntuales  $tr_t$ ,  $sn_t$ ,  $cl_t$  e  $ir_t$  de los factores apenas mencionados. El procedimiento empieza con el cálculo de las medias móviles centradas,  $CMA_t$ . La media móvil centrada se considera como una estimación de  $TR_t + CL_t$ . Puesto que el modelo

$$y_t = TR_t + SN_t + CL_t + IR_t$$

quiere decir que

$$SN_t + IR_t = y_t - (TR_t + CL_t)$$

se infiere entonces que la estimación  $sn_t + ir_t$  de  $SN_t + IR_t$  es

$$sn_t + ir_t = y_t - (tr_t + cl_t) = y_t - CMA_t$$

Con objeto de determinar  $sn_t$ , agrupamos los valores de  $sn_t + ir_t$  por estaciones (meses, trimestres, etc., como sea apropiado). Para cada estación, calculamos la media de los valores  $sn_t + ir_t$  de esa estación. Determinamos los factores estacionales normalizando los valores  $\bar{sn}_t$  de modo que los valores normalizados sumen cero. La normalización se consigue restando la cantidad  $\sum_{t=1}^L \bar{sn}_t / L$  de cada uno de los valores  $\bar{sn}_t$ . Es decir, la estimación de  $SN_t$  es

$$sn_t = \bar{sn}_t - \left( \sum_{t=1}^L \bar{sn}_t / L \right)$$



Enseguida calculamos que la observación compensada respecto a la variación estacional en el periodo  $t$  será

$$d_t = y_t - sn_t - \text{Variación estacional}$$

Al restar  $sn_t$  de la observación  $y_t$  se elimina la estacionalidad de los datos y podemos estimar mejor la tendencia. Estimamos  $tr_t$  de la tendencia TR, mediante el ajuste de una ecuación de regresión a los datos compensados respecto a la variación estacional. Por ejemplo, una tendencia lineal

$$TR_t = \beta_0 + \beta_1 t$$

o una tendencia cuadrática

$$TR_t = \beta_0 + \beta_1 t + \beta_2 t^2$$

se podría ajustar a las observaciones compensadas respecto a la variación estacional.

Puesto que el modelo

$$y_t = TR_t + SN_t + CL_t + IR_t$$

quiere decir que

$$CL_t + IR_t = y_t - TR_t - SN_t$$

se infiere que calculamos la estimación de  $CL_t + IR_t$  como

$$cl_t + ir_t = y_t - tr_t - sn_t$$

Con objeto de promediar  $ir_t$ , calculamos una media móvil de tres periodos de los valores  $cl_t + ir_t$ . Es decir, la estimación de  $CL_t$  es

$$cl_t = \frac{(cl_{t-1} + ir_{t-1}) + (cl_t + ir_t) + (cl_{t+1} + ir_{t+1})}{3}$$

Para terminar, calculamos la estimación de  $IR_t$ :

$$ir_t = (cl_t + ir_t) - cl_t$$

En general, las estimaciones  $tr_t$ ,  $sn_t$ ,  $cl_t$  e  $ir_t$  se usan para describir series temporales. También se utilizan para calcular predicciones. Si no hay patrón en el componente irregular, predecimos que  $IR_t$  es igual a cero. Entonces, el pronóstico puntual de  $y_t$  es

$$\hat{y}_t = tr_t + sn_t + cl_t$$

si existe un ciclo muy bien definido que se puede predecir. El pronóstico puntual es

$$\hat{y}_t = tr_t + sn_t$$

si no existe un ciclo muy bien definido o si  $CL_t$  no se puede predecir. Aunque no hay intervalo de predicción teóricamente correcto para  $y_t$ , un intervalo de predicción  $100(1 - \alpha)\%$  aproximado para  $y_t$  es

$$[\hat{y}_t \pm B_t[100(1 - \alpha)]]$$

donde  $B_t[100(1 - \alpha)]$  es el límite de error en un intervalo de predicción de  $100(1 - \alpha)\%$

$$[tr_t \pm B_t[100(1 - \alpha)]]$$

para la observación compensada respecto a la variación estacional  $d_t = y_t - sn_t$ .

## 7.3 MÉTODO DE AJUSTE ESTACIONAL X-12-ARIMA

Una de las principales actividades de la Oficina de Censos de Estados Unidos es efectuar un ajuste estacional, es decir, eliminar los componentes o efectos estacionales de los datos. Esta oficina proporciona tanto series temporales sin ajuste estacional como series temporales con ajuste estacional relacionadas con todos los aspectos económicos y comerciales. Entre los ejemplos se encuentran series temporales sobre el inicio de una construcción, ventas al menudeo y comercio con el extranjero. Además, proporciona a otros organismos, incluso a gobiernos extranjeros, programas para computadoras para efectuar ajustes estacionales. Por lo tanto, identificar los componentes estacionales de una serie temporal es una tarea importante. El último programa para computadora es X-12-ARIMA Seasonal Adjustment Program (Findley *et al.*, 1998). Éste es una extensión del programa para el ajuste estacional Statistics Canada's X-11 ARIMA (Dagum, 1980), el cual, a su vez, es una extensión de U.S. Census Bureau's X-11 (Shiskin, Young y Musgrave, 1967). Los métodos que siguen estos programas son mucho más complejos que los métodos de descomposición multiplicativa y aditiva de las secciones anteriores. Por ejemplo, los datos originales se ajustan a las "variaciones de los días hábiles para el comercio". Es decir, los datos se ajustan para explicar el hecho de que, por ejemplo, diferentes meses o trimestres constan de una cantidad distinta de días hábiles. Hay métodos automáticos para detectar las observaciones atípicas (observaciones inusuales). Los modelos de ARIMA se podrían usar para desplegar las series temporales mediante el pronóstico de valores futuros. Como se puede ver en la Tabla 7.2(a), los periodos se pierden tanto al inicio como al final de la serie temporal cuando se determinan las medias móviles. Se obtienen estimaciones más exactas para los componentes estacionales mediante los datos desplegados. Los modelos de ARIMA (autoregressive integrated moving average) se denominan modelos de Box-Jenkins en este texto y se estudian en los Capítulos 9 a 12.

Los últimos programas para ajuste estacional son muy complejos, uno de sus elementos es el método de la "razón de la media móvil", que es el que se aplica en el Ejemplo 7.1 cuando se calcula

$$\frac{y_t}{\text{CMA}_t} = \frac{y_t}{\text{tr}_t \times \text{cl}_t} = \text{sn}_t \times \text{ir}_t$$

Luego se obtienen los promedios de estos valores por mes. Estamos suponiendo que los componentes estacionales no cambian con el tiempo. En cambio, el procedimiento X-12-ARIMA efectúa un procedimiento de media móvil ponderada sobre los valores que corresponden al mismo mes. Este procedimiento de la media móvil elimina el componente irregular pero permite que el componente estacional que corresponde al mismo mes del año cambie con el tiempo. Por ejemplo, una elección en X-12-ARIMA es una media móvil estacional ponderada que se calcula como sigue:

$$\overline{\text{sn}}_t = \frac{1}{9} \text{sn}_{t-24}^{(1)} + \frac{2}{9} \text{sn}_{t-12}^{(1)} + \frac{3}{9} \text{sn}_t^{(1)} + \frac{2}{9} \text{sn}_{t+12}^{(1)} + \frac{1}{9} \text{sn}_{t+24}^{(1)}$$

donde

$$sn_t^{(1)} = \frac{y_t}{CMA_t} = \frac{y_t}{tr_t \times cl_t} = sn_t \times ir_t$$

Evidentemente, necesitamos datos de muchos años para esta media móvil estacional. Al igual a lo que se hizo en el método de descomposición clásico, es necesario normalizar estos promedios. El lector interesado puede referirse a Findley *et al.* (1998) donde encontrará más detalles sobre los procedimientos en X-12-ARIMA.

Entender los métodos de descomposición de las Secciones 7.1 y 7.2 redundará en importantes conocimientos para aprender más sobre métodos de descomposición. En muchos programas de pronósticos se proporcionan opciones para ajustar estacionalmente series temporales o para compensar series temporales respecto a la variación estacional. Es importante entender en cada caso la manera como se determina el componente estacional  $sn_t$  antes de que una observación  $y_t$  sea compensada respecto a la variación estacional calculando  $y_t/sn_t$ . Entre las preguntas obvias está si los “días hábiles” se consideran y cómo se detectan las observaciones atípicas y cómo se eliminan.