التمهيد الأسى Exponential Smoothing

طريقة المتوسطات المتحركة البسيطة السابق ذكرها من الطرق الخاصة جدًا والتي تفيد فقط إذا كانت بيانات السلسلة يغلب عليها الطابع العشوائي، أي إذا كانت البيانات تتأرجح بشكل غير نمطي حول متوسط ثابت يمثل مستوى السلسلة على الفترة الزمنية موضع الدراسة (والتر، 1992 صفحة 235). ولكن في الكثير من التطبيقات قد يتغير متوسط الظاهرة ببطء على الفترة الزمنية موضع الدراسة، وفي مثل هذه الحالات قد يكون من المنطقي إعطاء وزن أكبر لأحدث مشاهدة عند التنبؤ وأوزانًا تتناقص بشكلٍ أو بآخر بزيادة عمر المشاهدة أي بزيادة الفاصل الزمني بين زمن المشاهدة والزمن الذي يراد التنبؤ عنده. وبالطبع يوجد العديد من الدوال الرياضية التي تعكس مفهوم تناقص الأوزان أو الأهميات بزيادة عمر المشاهدة، إلاّ أن أهم هذه الدوال ما يعرف بالدوال الأسية والتي وجدت أرضية خصبة وممهدة ليس في أدبيات السلاسل الزمنية التقليدية فحسب بل في الأدبيات الحديثة أيضًا. وتعتمد فكرة هذه الدوال على إعطاء وزن ترجيح كبير لأحدث مشاهدة عند الزمن الذي يراد التنبؤ عنده ثم إعطاء أوزان ترجيحية تتناقص بشكل أسي مع زيادة الفاصل الزمني بين زمن التنبؤ وزمن المشاهدة.

إذا افترضنا أننا نقف عند نقطة أصل معينة t ونريد التنبؤ بقيمة الظاهرة عند الفترة الزمنية $t+1$ وأن هذا التنبؤ يرمز له بالرمز $\hat{y}_1(1)$, ، فإن نموذج التمهيد الأسي يعرف على الصورة:

$$\hat{y}_1(1) = \frac{1}{c}\left[y_t + wy_{t-1} + w^2y_{t-2} + \cdots w^{t-1}y_t\right] \qquad \ldots\ldots\ldots\ldots\ldots\ldots.(1.6.8)$$

حيث:

$$0 < w < 1 \;\; ; \;\; c = \sum_{i=0}^{t-1} w^t = \frac{(1-w^t)}{(1-w)}$$

ويسمي العدد w بمعامل التناقص discount coefficient، ويشير الثابت c ، إلى مجموع أوزان الترجيح التي تجعل من النموذج (1.6.8) متوسط حقيقي. ويعرف النموذج (1.6.8) عادة في أدبيات السلاسل الزمنية بنموذج المتوسطات المتحركة المرجح آسيًا exponentially weighted moving average model ويشار إليه عادة بالحروف EWMA. وتعتمد قيمة معامل التناقص w على سرعة التغير في مستوي السلسلة وعادة ما يكون $0.7 < w < 0.95$

إذا كانت t كبيرة فإن $w^t \to 0$ وفي هذه الحالة يمكن إعادة كتابة النموذج (1.6.8) على الصورة:

$$\hat{y}_t(1) = (1-w)y_t + (1-w)wy_{t-1} + (1-w)w^2y_{t-2} + \cdots .... \qquad (1.6.9)$$

وتجدر الإشارة إلى أن $\hat{y}_t(1)$ لا يمثل بالفعل متوسطًا حقيقيًا لأن مجموع أوزان الترجيح $(1-w), (1-w)w, (1-w)w^2 \ldots ..$ يساوي الواحد الصحيح. كما تجدر الإشارة إلى أن نظام الترجيح (1.6.9) يمكن تعديله باختيار قيم مختلفة للمعامل w ، فإذا كانت w كبيرة فإن وزن الترجيح الذي يعطى

المشـــاهدة الحالية لا يكون صـــغيرًا وأوزان الترجيح المتتابعة تتناقص ببطء، أما إذا كانت w صـــغيرة فإن وزن الترجيح الذي يعطى للمشاهدة الحالية يكون كبيرًا وأوزان الترجيح المتتابعة تتناقص بسرعة. فعلى سبيل المثال إذا كـانـت $w = 0.9$ فـإن معـامـلات المشـــاهـدات $y_t, y_{t-1}, y_{t-2}, \dots$ تـكون علـى الـترتـيـب $0.1, 0.09, 0.081, \dots$ أي أن تأثير المشـاهدات يتضـاءل ببطء، أما إذا كانت $w = 0.1$ فإن معاملات نفس المشاهدات $y_t, y_{t-1}, y_{t-2}, \dots$ لاستكون على الترتيب $0.9, 0.09, 0.009$ أي أن تأثير المشاهدات يتضاءل بسرعة بدءًا من المشاهدة الحالية.

ويمكن كتابة التنبؤ عند الفترة الزمنية السابقة $(t-1)$ من الصورة $(1.6.9)$ كما يلي:

$$\hat{y}_{t-1}(1) = (1-w)y_{t-1} + (1-w)wy_{t-2} + (1-w)w^2y_{t-3} + \cdots \quad (1.6.10)$$

وبالتعويض من $(1.6.10)$ في $(1.6.9)$ نصـل إلى الصـيغة التكرارية (التتابعية) Recurrence Formula الآتية:

$$\hat{y}_t(l) = (1-w)y_t + w\hat{y}_{t-I}(I) \qquad (1.6.11)$$

وتوضـح الصيغة $(1.6.11)$ أن التنبؤ الحديث عند الزمن t يساوي الوسط الحسابي المرجح للتنبؤ السابق له مباشرة $\hat{y}_{t-I}(I)$ والمشاهدة الحديثة y_t ووزني الترجيح هما w و $(1-w)$ على الترتيب وتستأثر المشاهدة الأحدث بالنصيب الأكبر في هذه العلاقة إذا كانت قيمة w صـغيرة، بينما تكون مساهمة هذه المشاهدة صـغيرة إذا كانت قيمة w كبيرة. وتمكن الصيغة $(1.6.11)$ من حساب التنبؤات للسلسلة بشكل تتابعي كالآتي:

$$\hat{y}_1(I) = (1-w)y_1 + w\hat{y}_0(1) \qquad (1)$$
$$\hat{y}_2(I) = (1-w)y_2 + w\hat{y}_1(1) \qquad (2)$$
$$\hat{y}_3(1) = (1-w)y_3 + w\hat{y}_2(1) \qquad (3)$$
$$\vdots$$

بالتعويض من المعادلة (1) في المعادلة (2):

$$\hat{y}_2(1) = (1-w)y_2 + w(1-w)y_1 + w^2\hat{y}_0(1) \qquad (4)$$

بالتعويض من المعادلة (4) في المعادلة (3):

$$\hat{y}_3(1) = (1-w)y_3 + w(1-w)y_2 + w^2(1-w)y_1 + w^3\hat{y}_0(1)$$

بالاستمرار في هذه العملية نصل إلى:

$$\hat{y}_t(1) = [(1-w)y_t + w(1-w)y_{t-1} + w^2(1-w)y_{t-2} + \cdots + w^{t-1}y_1] + w^1\hat{y}_0(1) \qquad (1.6.12)$$

وتوضـح الصـورة $(1.6.12)$ أن التنبؤ عند الفترة الزمنية t، يمكن التعبير عنه بدلالة مشـاهدات السـلسـلة المتاحة $y_t, y_{t-1}, \dots . y_1$ والتنبؤ الابتدائي $\hat{y}_0(1)$. وتجدر الإشارة هنا إلى ملاحظتين أساسيتين حول الصيغة $(1.6.12)$. الملاحظـة الأولى أن معـامـلات شـــاهـدات $y_t, y_{t-1}, \dots . y_1$ هـــــي على الترتيـب $(1-w), (1-w)w, (1-w)w^2, \dots (1-w)w^{t-1}$ وأن هـذه المعـاملات تقيس مسـاهمـة هـذه المشـــاهدات في التنبؤ $\hat{y}_t(1)$. وكما هو واضـــح أن هذه المعاملات تعطي أهميات أكبر للمشـــاهدات الحديثة

وتتناقص قيم هذه المعاملات بشـــكل أســـي. والملاحظة الثانية أن المعامل w^t يقيس مســـاهمة التنبؤ المبدئي $\hat{y}_0(1)$ على التنبؤ $\hat{y}_1(1)$ ، ومن ثم يتضـــــاءل تأثير التنبؤ المبدئي $\hat{y}_0(1)$ على التنبؤ $\hat{y}_1(1)$ إذا كان طول السلسلة n معقولاً. ولحساب التنبؤات عادة ما تستخدم الصيغة التكرارية (1.6.11) لسهولة تحديث التنبؤات حيث يكفي معرفة المشـــاهدة الحديثة والتنبؤ الســـابق مباشـــرة. ولحســـاب التنبؤات يجب معرفة القيمة الابتدائية $\hat{y}_0(1)$ ومعامل التناقص w . بالنسبة لاختيار القيمة الابتدائية يوجد العديد من الطرق لتقديرها أهمها:

1- استخدام الوسط الحسابي لقيم السلسلة $y_1, y_2, \dots, y_n$، وهذه الطريقة عادة ما تكون ملائمة إذا كان متوسط السلسلة يتغير ببطء على الفترة الزمنية موضع الدراسة.

2-ويفضل بعض الباحثين استخدام المشاهدة الأولي y_1 كتقدير للقيمة الابتدائية.

3- استخدام الوسط الحسابي لبعض المشاهدات الأولى لتقدير هذه القيمة.

للمزيد من التفاصـــــيل حول اختيار القيمة الابتدائية يمكن للقارئ الرجوع إلى (1962) Brown أو إلى (1976) Montgomery and Johnson أو إلى (1983) Abraham and Ledolter أو إلى (1987) Bowerman and O'Connell

ويســـــتأثر معامل التناقص w بأهمية خاصـــــة في التأثير على التنبؤ $\hat{y}_t(1)$, ومن ثم يحظى هذا المعامل بأهمية خاصة عند اختياره. وهناك بعض الخطوط العامة التي يمكن الاسترشاد بها عند هذا الاختيار والخاصة بسرعة التقلبات التي تحدث في السلسلة. فإذا كانت السلسلة تتعرض للكثير من التقلبات غير المنتظمة فقد يكون من الأفضل استخدام قيمة كبيرة للمعامل w وذلك من أجل إعطاء وزن وأهمية للتنبؤ السابق $\hat{y}_{t-1}(1)$, أكبر من وزن المشاهدة الحديثة $\hat{y}_t(1)$، ويؤدي هذا إلى التخلص من الأخطاء العشوائية والحصول على تنبؤات مستقرة. أما إذا كانت السلسلة أكثر هدوءًا واستقرارًا أو يوجد تغير منتظم في نمط السلسلة فقد يكون من الأفضل اختيار قيمة صـــغيرة للمعامل w وذلك من أجل إعطاء وزن أكبر للمشـــاهدة الحديثة (شـــعراوي، 2005 صـــفحة 31). ولكن هذه الخطوط العامة لا تمكن بالطبع من اختيار قيمة دقيقة لهذا المعامل في التطبيقات العملية ولذلك عادة ما يتم اختيار هذا المعامل عن طريق المحاكاة simulation في مثل هذه التطبيقات حيث يتم توليد تنبؤات مختلفة للمعامل w ثم تقارن هذه التنبؤات بالقيم العملية للسلسلة $y_1, y_2, \dots, y_n$, لحساب الأخطاء المناظرة لكل قيمة من قيم المعامل w. بعد ذلك يتم حساب أحد المقاييس التي سبق دراستها لقياس حجم الأخطاء وليكن مجموع المربعات المناظر لكل قيمة من قيم w ، وتكون قيمة المعامل w المناســـبة هي القيمة التي تجعل هذا المجموع أقل ما يمكن. وطريقة المحاكاة ليســـت الطريقة الوحيدة المعروفة لاختيار قيمة w ولكن هناك طرق أخرى أهمها ما يعرف بطريقة التجربة والخطأ trial and error ، ولن نتعرض لدراسة هذه الطريقة هنا ولكن يمكن للقارئ الرجوع إلى (1994) Gaynor and Kirkpatrick للتعرف عليها والمثال الآتي يوضح كيفية استخدام طريقة التمهيد الأسي في التنبؤ .

مثال (4): البيانات الآتية تمثل عدد الأجهزة (بالمائة) المباعة التي ســـجلت شـــهريًا في دفاتر إحدى الشركات:

14 16 17 13 15 14 15 13 14 12 12 11 16

قدر القيمة الابتدائية $\hat{y}_0(1)$ باستخدام الوسط الحسابي لقيم السلسلة ثم استخدم هذه القيمة لإيجاد التنبؤات المناظرة مرة باستخدام $w = 0.5$ ومرة أخرى باستخدام $w = 0.9$، أي التنبؤات أفضل؟ اشرح سبب إجابتك.

الحـل:

$$\hat{y}_0(1) = \bar{y} = \frac{1}{13}[11 + 12 + 12 + \cdots + 16] = 14$$
$$\hat{y}_t(1) = (1 - w)y_t + \mathrm{w}\hat{y}_{t-1}(1)$$

إذا كانت $w = 0.7$:

$$\hat{y}_1(1) = 0.3(y_1) + 0.7\hat{y}_0(1) = 0.3(11) + 0.7(14) = 13.1$$
$$\hat{y}_2(1) = 0.3(y_2) + 0.7\hat{y}_1(1) = 0.3(12) + 0.7(13.1) = 12.77$$
$$\hat{y}_3(1) = 0.3(y_3) + 0.7\hat{y}_2(1) = 0.3(12) + 0.7(12.77) = 12.539$$

وبالاستمرار في هذه العملية يمكن توليد التنبؤات الموضحة في جدول (1)

إذا كانت $w = 0.9$:

$$\hat{y}_1(1) = 0.1(y_1) + 0.9\hat{y}_0(1) = 0.1(11) + 0.9(14) = 13.7$$
$$\hat{y}_2(1) = 0.1(y_2) + 0.9\hat{y}_1(1) = 0.1(12) + 0.9(13.7) = 12.53$$
$$\hat{y}_3(1) = 0.1(y_3) + 0.9\hat{y}_2(1) = 0.1(12) + 0.9(13.53) = 13.377$$

وهكذا يمكن توليد باقي التنبؤات والموضحة في جدول (1)

جدول (1): القيم الفعلية والتنبؤات لبيانات المثال (4)

t	y_t	التنبـــــؤات		$e^2{}_t = [y_t - \hat{y}_{t-1} - 1)]^2$	
		$w = 0.7$	$w = 0.9$	$w = 0.7$	$w = 0.9$
1	11	14	14	9.00	9.00
2	12	13.1	13.7	1.21	2.89
3	12	12.77	13.53	0.59	2.34
4	14	12.539	13.38	2.13	0.38
5	13	12.98	13.34	0.00	1.12
6	15	12.98	13.51	4.08	2.22
7	14	13.59	13.66	0.17	0.12
8	15	13.17	13.69	3.35	1.72

9	13	14.10	13.82	1.21	0.67
10	17	13.77	13.74	10.43	10.63
11	16	17.74	14.07	1.59	3.72
12	14	15.12	14.26	1.25	0.07
13	16	14.78	14.23	1.49	3.13
		15.15	14.41	–	–

بفحص نتائج الجدول (1) نجد أن مجموع مربعات الأخطاء المناظرة للقيمة $w = 0.7$ هو:

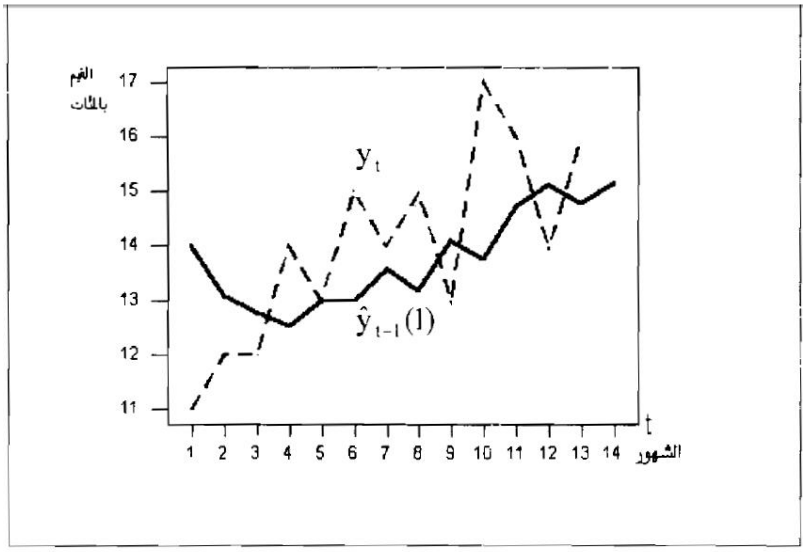
$$S\,(0.7) = 9 + 1.121\ + 0.59 + \ldots + 1.4\,9 = 35.5$$

وبالمثل فإن مجموع مربعات الأخطاء المناظرة للقيمة $w = 0.9$ هو:

$$S(0.9) = 9 + 2.89\ + 2.34 + \ldots + 3.13 = 37.01$$

حيث إن $S\,(0.7) < S(0.9)$ يمكن القول بأن التنبؤات المناظرة لقيمة $w = 0.7$ أفضـــل من التنبؤات المناظرة لقيمة $w = 0.9$، ويكون التنبؤ بعدد الأجهزة المباعة في الشهر التالي هو 1515 جهازا ويعرض شكل (7) القيم الفعلية والتنبؤات التي حصلنا عليها إذا كانت $w = 0.7$

شكل (7): القيم الفعلية والتنبؤات للمثال (4)



ويلاحظ من شكل (7) أن سلسلة التنبؤات أكثر تمهيدًا ونعومة من السلسلة الأصلية.

وقد أوجد نموذج المتوســطات المتحركة المرجح أســيًا لنفســه طريقا ممهدًا في الكثير من تطبيقات وأدبيات التنبؤ في مجالات المعرفة المختلفة خاصة الاقتصادية والإدارية منها وذلك لعدة أسباب، السبب الأول هو سهولة تطبيقه وتحديث تنبؤاته باستخدام التنبؤ السـابق والمشـاهدة الأحدث فقط، وهذا السـبب هو أحد الأسـباب الهامة

لانتشار طريقة التمهيد الأسي في هذه المجالات خاصة عند التنبؤ بالعديد من السلاسل الزمنية كما هو الحال عند التنبؤ بمبيعات الآلاف من السلع المختلفة الموجودة في أحد المحلات العملاقة. السبب الثاني في هذا الانتشار يعود إلى آلية الطريقة الكاملة، فبمجرد برمجة الخطوات الرئيسية واختيار قيمة المعامل w يمكن الحصول على التنبؤات دون تدخل الإرادة الإنسانية، هذه الميزة بالذات هامة جدا لمستخدمي الإحصاء في مجالات المعرفة المختلفة والذين يبحثون دائما عن أساليب لا تتطلب مهارات وخبرات خاصة. السبب الثالث وراء انتشار هذه الطريقة هو سهولة برمجة خطواتها الرئيسية وانخفاض تكاليف استخدامها مقارنة باستخدام النماذج العشوائية الحديثة. السبب الرابع أن استخدام هذه الطريقة لا يتطلب توافر عدد كبير من المشاهدات ولذلك وجدت هذه الطريقة أرضًا خصبة في البلاد النامية حيث تعاني معظم هذه البلاد من عدم توافر الحد الأدنى من المشاهدات الضروري لاستخدام النماذج العشوائية الحديثة.

وتتعرض طريقة التمهيد الأسي للعديد من الانتقادات، أولها عدم وجود منهجية عامة للاختيار بين أنظمة الترجيح البديلة وعدم وجود طريقة عامة لتقويم نظام الترجيح المختار. والانتقاد الثاني الذي يوجه إلى هذه الطريقة أنها تعالج كل السلاسل الزمنية التي تنشأ في الواقع بطريقة واحدة ولذلك فهي قد تؤدي إلى تنبؤات غير صالحة، والانتقاد الثالث الذي يوجه إلى هذه الطريقة أنه ليس من السهل دائما اختيار المعامل w بدقة وعدم وجود طريقة وحيدة لاختيار هذا المعامل والتقدير المبدئي $\hat{y}_0(1)$ ، ولذلك فإن التنبؤات التي نحصل عليها من هذه الطريقة قد تختلف من باحث إلى آخر. الانتقاد الأخير والهام أن هذه الطريقة تعتبر من الطرق الخاصة وذلك لأنها تكون ملائمة فقط إذا كانت العملية العشوائية الكامنة في البيانات لها مواصفات خاصة. بصورة أكثر تحديدًا يمكن القول بأن هذه الطريقة تعمل بصورة جيدة إذا كانت العملية العشوائية تنتمي إلى فئة جزئية فقط من النماذج العشوائية الحديثة والتي لها مواصفات خاصة والتي تكون مفيدة في بعض المواقف وغير المفيدة في مواقف أكثر (شعراوي، 2005 صفحة 39).