المحاضرة الثالثة (03)

الأهداف: في نهاية هذه المحاضرة يكون الطالب قادرا على:

- التنبؤ باستخدام نموذج التمهيد الأسي.
- معرفة مميزات وعيوب النموذج.

التمهيد الأسي (Exponential Smoothing)

طريقة المتوسطات المتحركة البسيطة السابق ذكرها من الطرق الخاصة جدا والتي تفيد فقط إذا كانت بيانات السلسلة يغلب عليها الطابع العشوائي أي إذا كانت البيانات تتأرجح بشكل عشوائي حول المتوسط ثابت يمثل مستوى السلسلة خلال الفترة موضع الدراسة. ولكن في كثير من التطبيقات قد يتغير متوسط الظاهرة ببطء على الفترة الزمنية موضع الدراسة، وفي هذه الحالات قد يكون من المنطقي إعطاء وزن اكبر لأحدث مشاهدة عند التنبؤ وأوزانا تتناقص بشكل 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 اوزانها أو الأهميات بزيادة عمر المشاهدة، إلا أن اهم هذه الدوال ما يعرف بالدوال الأسية والتي وجدت أرضية خصبة وممهدة ليس في أدبيات السلاسل الزمنية التقليدية فحسب بل في الأدبيات الحديثة أيضا.

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.

صيغة النموذج الرياضية

إذا افترضنا أننا نقف عند نقطة زمنية معينة t ونريد التنبؤ بقيمة الظاهرة عند الفترة الزمنية t+1 وان هذا التنبؤ يرمز له بالرمز $\hat{y}_{t+1}$. فإن نموذج التمهيد الأسي يعرف بالصورة التالية:

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

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

ويسمى العدد w بمعامل التناقص ويشير الثابت c إلى مجموع الأوزان الترجيح التي تجعل من النموذج (1) متوسط حقيقي، ويعرف النموذج (1) عادة في أدبيات السلاسل الزمنية بنموذج المتوسطات المتحركة المرجحة أسيا (*Exponentially Weighted Moving Average Model*) ويشار إليه عادة بالرموز EWMA. وتعتمد قيمة معامل التناقص W على سرعة التغير في مستوى السلسلة عادة ما يكون: $0.7 < w < 0.95$

إذا كانت t كبيرة فإن: $w^t \to 0$ وفي هذه الحالة يمكن إعادة كتابة النموذج (1) على الصورة التالية:

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

وتجدر الإشارة إلى أن $\hat{y}_{t+1}$ يمثل بالفعل متوسطا حقيقيا لأن مجموع أوزان الترجيح

$(1-w), (1-w)w, (1-w)w^2y_{t-2}, \ldots = 1$ يساوي الواحد الصحيح. كما تجدر الإشارة على أن نظام الترجيح (2) يمكن تعديله باختيار قيم مختلفة للمعامل w. فإذا كانت w كبيرة فإن وزن الترجيح الذي يعطي للمشاهدة الحالية y_t يكون صغيرا وأوزان الترجيح المتابعة تتناقص ببطء. أما إذا كانت w صغيرة فإن وزن الترجيح الذي يعطي للمشاهدة الحالية يكون كبيرا وأوزان الترجيح المتابعة تتناقص بسرعة.

ويمكن كتابة التنبؤ عند الفترة الزمنية (t) من الصورة الرياضية (2) كما يلي:

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

وبتعويض المعادلة (3) في المعادلة رقم (2) نصل إلى الصيغة التكرارية (التتابعية) الآتية:

$$\hat{y}_{t+1} = (1-w)y_t + w\hat{y}_t \ldots . (4)$$

وتوضح الصيغة (3) أن التنبؤ الحديث عند الزمن $t+1$ يساوي الوسط الحسابي المرجح للتنبؤ السابق له مباشرة $\hat{y}_t$ 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 صغيرة. بينما تكون مساهمة هذه المشاهدة صغيرة إذا كانت قيمة w كبيرة. وتمكن الصيغة الرياضية (4) من حساب التنبؤات للسلسلة بشكل تتابعي كالاتي:

$$\hat{y}_2 = (1-w)y_1 + w\hat{y}_1$$

$$\hat{y}_3 = (1-w)y_2 + w\hat{y}_2$$

$$\hat{y}_4 = (1-w)y_3 + w\hat{y}_3$$

.

.

.

ولحساب التنبؤات نستخدم الصيغة التكرارية لسهولة تحديث التنبؤات حيث يكفي معرفة المشاهدة الحديثة والتنبؤ السابق مباشرة. ولحساب التنبؤات يجب معرفة القيمة الابتدائية $\hat{y}_1$ ومعامل التناقص w.

بالنسبة لاختيار القيمة الابتدائية يوجد العديد من الطرق لتقديرها وأهمها:

1. استخدام الوسط الحسابي لقيم السلسلة $y_1, y_2, \dots\dots, y_n$ وهذه الطريقة عادة ما تكون ملاءمة إذا كان متوسط السلسلة يتغير ببطء على الفترة الزمنية موضع الدراسة. (وهذه هي الطريقة التي سنعتمدها).
2. يفضل بعض الباحثين استخدام المشاهدة الأولى y_1 كتقدير للقيمة الابتدائية $\hat{y}_1$.
3. استخدام الوسط الحسابي لبعض المشاهدات الأولى لتقدير القيمة.

ويستأثر معامل التناقص w بأهمية خاصة 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 أكبر من وزن المشاهدة الحديثة. وهذا يؤدي إلى التقليل من الأخطاء العشوائية والحصول على تنبؤات مستقرة.

أما إذا كانت السلسلة أكثر هدوء واستقرارا أو يوجد تغير منتظم في نمط السلسلة فقد يكون من الأفضل اختيار قيمة صغيرة للمعامل w وذلك من أجل إعطاء وزن أكبر للمشاهدة الحديثة.

ولكن هذه الخطوط العامة لا تمكن بالطبع من اختيار قيمة دقيقة لهذا المعامل في التطبيقات الحديثة العملية، ولذلك عادة ما يتم اختيار هذا المعامل من خلال التجربة والخطأ حيث يتم توليد تنبؤات مختلفة للمعامل w ثم تقارن هذه التنبؤات بالقيم العملية أو الحقيقية للسلسلة الزمنية لحساب الأخطاء المناظرة لكل قيمة من قيم المعامل w. بعد ذلك يتم حساب أحد المقاييس التي سبق دراستها لقياس حجم الأخطاء وليكن متوسط مجموع مربعات الأخطاء المنظر لكل قيمة من قيم w ، وتكون قيمة المعامل w المناسبة هي القيمة التي تجعل هذا المتوسط لمجموع مربعات الأخطاء أقل ما يمكن.

الانتقادات

تتعرض طريقة التمهيد الأسي للعديد من الانتقادات، أهمها:

1. عدم وجود منهجية عامة للاختيار بين أنظمة الترجيح البديلة وعدم وجود طريقة عامة لتقويم نظام الترجيح المختار.
2. هذه الطريقة تعالج كل السلاسل الزمنية التي تنشأ في الواقع بطريقة واحدة وبالتالي قد تؤدي إلى تنبؤات غير صالحة.
3. ليس من السهولة اختيار المعامل المرجح او الموزون w بدقة.
4. عدم وجود طريقة وحيدة لتقدير القيمة الابتدائية المقدرة $\hat{y}_1$ ولذلك فغن التنبؤات التي نحصل عليها من هذه الطريقة قد تختلف من باحث لآخر.
5. هذه الطريقة تعمل بشكل جيد اذا كانت العملية العشوائية تنتمي على فئة جزئية فقط من النماذج العشوائية الحديثة والتي قد تكون مفيدة في بعض الحالات وغير مفيدة في حالات أخرى اكثر.

مثال تطبيقي (انظر التمرين الثالث من السلسلة رقم 01)

ليكن لديك الجدول التالي الذي يوضح عدد الأجهزة المباعة (بالمئة) التي سجلت شهريا في دفاتر إحدى الشركات

الشهر	1	2	3	4	5	6	7	8	9	10	11	12	13
النسبة	11	12	12	14	13	15	14	15	13	17	16	14	16

- قدر القيمة الابتدائية $\widehat{y_0}$ باستخدام الوسط الحسابي لقيم السلسلة ثم استخدم هذه القيمة لإيجاد التنبؤات المناظرة مرة باستخدام w=0.7 ومرة باستخدام w=0.9 أي التنبؤات أفضل؟ برر اجابتك؟

الإجابة

- تقدير القيمة الابتدائية $\widehat{y_1}$ باستخدام الوسط الحسابي لقيم السلسلة
- $\widehat{y_1} = \bar{y} = \frac{\sum y}{n} = \frac{11+12+\cdots.16}{13} = 14$
- استخدام القيمة الابتدائية لإيجاد التنبؤات المناظرة مرة باستخدام w=0.7 ومرة باستخدام w=0.9

بالتعويض في معادلة التمهيد الأسي التالية:

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

نتحصل على النتائج وفق الجدول التالي

مربعات الأخطاء		التنبؤات داخل السلسلة			
w = 0,9	w = 0,7	w = 0,9	W = 0,7	عدد الأجهزة المباعة بالنسبة المئوية	الأشهر
9	9	14	14	11	1
2,89	1,21	13,70	13,10	12	2
2,34	0,59	13,53	12,77	12	3
0,39	2,13	13,38	12,54	14	4
0,19	0,00	13,44	12,98	13	5
2,57	4,06	13,40	12,98	15	6
0,20	0,17	13,56	13,59	14	7
1,96	1,66	13,60	13,71	15	8
0,55	1,21	13,74	14,10	13	9
11,11	10,44	13,67	13,77	17	10
4,00	1,59	14,00	14,74	16	11
0,04	1,25	14,20	15,12	14	12
3,31	1,48	14,18	14,78	16	13
				?	14
38,56	34,80	المجموع			
2,97	**2,68**	متوسط مربعات الأخطاء			

- للتنبؤ بنسبة الأجهزة المباعة في الشهر الرابع عشر 14 نستخدم الوزن الترجيحي $w = 0.7$ لأن هذا الوزن يناظر القيمة الأقل لمتوسط مربعات الأخطاء لنموذج التمهيد الأسي وبالتالي تكون القيمة التنبئية $\hat{y}_{14}$ كما يلي:

$$\hat{y}_{14} = (1 - 0.7)16 + (0.7)(14.78) = 15.15\%$$