إذا كان النموذج ضـربياً و $I_t = Y_t - T_t - S_t - C_t$ إذا كـان جمعيـاً. لاحـظ أن الضرب في ١٠٠ مرتين لإلغاء تأثير الضرب في ١٠٠ عن حساب S_t و C_t.

٢,٢,١٠ اختبارات لتقييم نجاح التجزئة :

بعد تطبيق طريقة التجزئة على السلسلة الزمنية وتقدير المكونات المختلفـة قـد تتساءل عما إذا كانت عملية التجزئة ناجحة. هناك عدة اختبارات تستخدم في الإجابة على مثل هذه الأسئلة نذكر بإيجاز بعضها.

٢,٢,١٠,١ اختبار الشهر المجاور Adjacent month test

هذا الاختبار مفيد بصفة خاصـة في السلاسـل الشـهرية عنـدما نقـدر الـدليل الموسمي S ونستخدمه لإزالة التأثير الموسمي من السلسلة ونرغب في معرفة ما إذا كان التأثير الموسمي قد أزيل من السلسلة. في هذه الحالة نحسب النسبة بين قيمة كـل شـهر ومتوسط قيمتي الشهر الذي يسبقه والذي يليه فى السلسلة الزمنية الأصلية. ثم نحسب هذه النسب للسلسلة بعد إزالة التأثير الموسمي فإذا كان هناك تأثيراً موسمياً في السلسلة الأصلية ونجحنا في إزالته فسنجد أن الاختلافات بين النسب في السلسلة الأصلية كبيرة بينما في السلسلة التي أزيل منها التأثير الموسمي صغيرة.

من ناحية أخري إذا حسبنا متوسط النسب لكل شهر (عبر السنوات) نحصل على صورة أوضح للاختلافات بين الشهور.

٢,٢,١٠,٢ اختبار يناير January test

إذا قسمنا كل قيمة في السلسلة التي أزيل منها التأثير الموسمي علـى القيمـة في يناير السابق نحصل على قيم معيارية يمثل شهر يناير فيها شهر الأساس. فإذا ظهر نمـط معين في هذه النسب فهذا يعني أن التأثير الموسمي لم تتم إزالته بشكل كامل. نلاحظ أن اختبار يناير يساعد في كشف أي موسمية داخل السنة بينما اختبار الشهر المجاور يكشف وجود الموسمية بين السنوات.

٣,١٠,٢,٢ اختبارات التغير المئوية **Percentage change test**

وتقوم على حساب التغير الذي حدث في أي شهر كنسبة مئوية من القيمة في الشهر السابق. فإذا كانت القيمة في شهر معين ١٢٠ وفي الشهر التالي له ١٣٠ فإن النسبة المئوية تكون

$$\left(\frac{130-120}{120}\right)\times 100 = 8.33\%$$

ويمكن إجراء هذا الاختبار على السلسلة الأصلية وعلى أساس كل من السلسلة الخالية من التأثير الموسمي ، سلسلة المكون العشوائي وسلسلة مكون الاتجاه العام والدوري (معاً).

مقارنة نتيجة اختبار التغير المئوي للسلسلة الخالية من التأثير الموسمي مع النتيجة المتحصل عليها من تطبيقه على السلسلة الأصلية يساعد في كشف حجم التغيرات الناتجة عن التأثير الموسمي. فإذا كان المتوسط الكلي لنسب التغير الشهري في السلسلة الأصلية ١٠,٥ مثلاً و كان المتوسط الكلي لنسب التغير الشهري في السلسلة الخالية من التأثير الموسمي ٢,٣ فإن نسبة التغير الشهري الناتجة عن التأثير الموسمي تكون $10.5 - 2.3 = 8.2$.

وإذا حسبت نسب التغير الشهري للسلسلة التي قيمتها المكونات غير المنتظمة (أو العشوائية) وكان متوسطها الكلي ١,٦ مثلاً فإن هذا الرقم يعطي مؤشراً للتغير الشهري في السلسلة الناتج عن التغيرات العشوائية، وواضح أن الفرق

$$10.5 - 8.2 - 1.6 = 0.7\%$$

يمكن إرجاعه للتغير الناتج عن **T** و **C** . لاحظ أن المتوسط الكلي للتغير في المكون العشوائي يمثل الحد الادني لخطأ التنبؤ المتوقع من السلسلة.

من ناحية أخري فإن تطبيق الاختبار على سلسلة الاتجاه العام – الدوري يبرز التغير الشهري فيها.

٢,٢,١١ طرق تجزئة أخري

منذ خمسينات القرن العشرين ظهر عدد من طرق التجزئة المطورة والـتي تمثـل جميعها طرقاً مطورة من طريقة التجزئة التقليدية. وتهدف هذه الطرق – والـتي تعـرف أيضا بطرق التعديل **adjustment methods** إلى تحسين مقدرات التأثير الموسمي ، الدوري ، الاتجاه العام والغير منتظم ، ومن ثم تعـديل السلسلة الزمنيـة بحيـث تغـدو خالية من التأثير الموسمي والغير منتظم لتمكين إبراز الاتجاه العام والـتغيرات الدوريـة فقط. وهي تقوم بذلك من خلال سلسلة من التعديلات والمتوسطات المتحركة (المعقدة أحياناً) مستفيدة من التطور الكبير في الحاسبات الآلية .

أولي هذه الطرق وأهمها هي الطريقـة المسـماة " تعـداد ٢ **Census II** " الـتي ابتكرها واستخدمها المكتب الأمريكـي للتعـداد عـام ١٩٥٥. الملامـح الرئيسـية لهـذه الطريقة عند تطبيقها على بيانات شهرية كما يلي :

١. تحسب النسب للمتوسط المتحرك كما في طريقة التجزئة التقليدية.

٢. يتم التعويض عن المتوسطات المتحركة التي تفقد في أول السلسلة وآخرها بتقديرات معينة.

٣. يقضى على التغيرات غير المنتظمـة في النسـب للمتوسـط المتحـرك بأخـذ متوسـط متحرك (مركب).

٤. تعدل النسب المعدلة بحيث يصبح مجموعها في كل سنة مساوياً ل١٢٠٠. تمثل هـذه النسب الآن تقدير مبدئي للعوامل الموسمية للأشهر المختلفة.

٥. تخلص السلسلة من التأثير الموسمي باستخدام العوامل الموسمية بخطـوة (٤) يعتـبر هذا تخليص أولي للسلسلة الزمنية من التأثير الموسمي.

٦. تطبيق متوسطات متحركة على السلسلة بخطوة (٥) للقضاء على أي آثار موسميـة وغير منتظمة لم يتم القضاء عليها بعد. ويتحقق ذلك من خلال سلسلة مـن الخطـوات المشابهة لتلك التي استخدمت للحصول على السلسلة الخالية (شكل أولى) من التـأثير الموسمي بخطوة (٥). أي أن السلسلة الخالية – أولياً– مـن التـأثير المـوسمي تسـتخدم

كنقطة بداية فتحسب متوسطات متحركة ، نسب لمتوسط متحرك ، تخليص من الأثر الموسمي ثم العشوائي للحصول على سلسلة خالية نهائياً من التأثير الموسمي.

وقد أدت الأبحاث المكثفة الموجهة نحو تحسين طرق تعديل السلسلة الزمنية إلى ظهور مجموعة الطرق المشار إليها بطرق X ومن أهمها $X-11-ARIMA$ و$X-12-ARIMA$. وبينما تستخدم census II مجموعة من المتوسطات المتحركة المتنوعة لتحسين تقدير القيم الضائعة (بسبب استخدام المتوسطات المتحركة) ولتحسين مقدرات التأثيرات الموسمية والتأثيرات الأخرى تلجأ $X-11-ARIMA$ و$X-12-ARIMA$ لاستخدام نماذج أريما للتنبؤ بالقيم التي تضيع في أول السلسلة الزمنية وآخرها. وتسبق طريقة $X-12-ARIMA$ عادة استخدام طريقة أخرى هي **Reg ARIMA** والتي تستخدم نموذج انحدار يمثل العلاقة بين قيم السلسلة ومتغيرات تمثل التقويم (أيام الشهر) بهدف معرفة أثر الاختلاف فيه وتعديل السلسلة في ضوء ذلك إضافة لمعرفة القيم الشاذة في السلسلة.

وتتوفر بعض البرمجيات الخاصة لتنفيذ تعديل السلاسل الزمنية. ومن بين هذه البرنامج **DEMETRA** والذي يستخدم لتنفيذ **X- ١١-ARIMA** .

الباب الثالث

التحليل الطيفى

٣,١ مقدمة

تفترض طريقة التجزئة أن قيم السلسلة الزمنية ناتجة عن أربعة أنواع من التأثيرات التي تعطى مفعولها بشكل جمعي أو ضربي ، والهدف من استخدام تلك الطريقة هو قياس كل من هذه التأثيرات أو عزلة. أما في التحليل الطيفى فينظر للسلسلة الزمنية على أنها نتاج لموجات جيب خفية والهدف من التحليل التوصل للموجات ذات التأثير الأكبر على السلسلة ومعرفة أطوالها وبالتالي عدد المرات التي تتكرر بها في مدى البيانات.

وفي هذا الباب نتعرف على أساسيات التحليل الطيفى.

٣,٢ دالة الجيب Sine function

تأخذ دالة الجيب الشكل

$$Y = \sin\theta$$

حيث سنفترض ، للتبسيط أن $0 \le \theta \le 2\pi$. وهي تكمل دورة كاملة (موجة) عندما تأخذ θ قيما بين 0 و 360^0 أو بين ٠ و 2π بمقياس الراديان.

وهناك ثلاثة خصائص لموجة الجيب وهي :

١. طول الموجة **Wave length**

ويقاس بطول المسافة بين أي قمتين متجاورتين تقعان في نفس الجانب من المحور الأفقي.

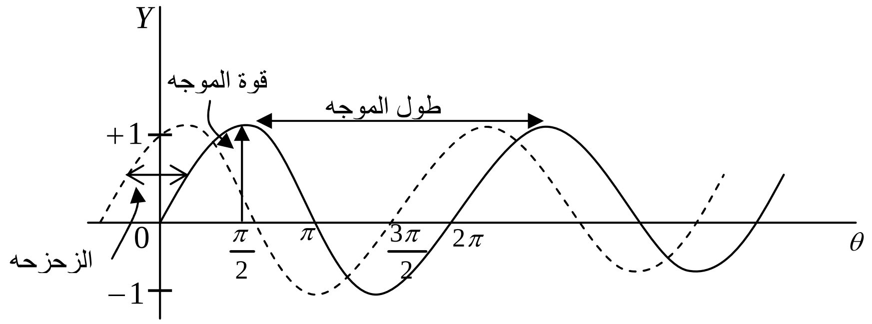
٢. قوة الموجة **Amplitude**

وهو أقصي ارتفاع للموجة ويعبر عن قوتها.

٣. الزحزحة **Shift**

ويقصد بها الزحزحة الأفقية للموجة عندما لا تبدأ من نقطة الأصل.

ويوضح شكل (٣,١) المقصود بكل من هذه المصطلحات



شكل (٣.١)

ويمكن إثبات أنه لأى سلسلة زمنية بحجم **n** المسافات الزمنية فيها متساوية يمكن دائماً تجزئها لموجات جيب بأطوال وتكرارات مختلفة. لكن إذا كانت **n** عدد فردي فإن عدد موجات الجيب التي يمكن توفيقها بطريقة المربعات الصغرى لا يزيد عن $(n-1)/2$ بينما العدد الأقصى الذي يمكن توفيقه في حالة عدد زوجي $(n-1)/2$.

٣,٢,١ توفيق دالة جيب واحدة بتكرار معروف

أفرض أن لدينا سلسلة زمنية حجمها **n** ونريد أن نوفق دالة جيب بموجة ذات تكرار f' عليها كما فعلنا حين سعينا لتوفيق خط مستقيم يمثل الاتجاه العام للسلسلة. نلاحظ أولاً أننا يجب أن نضع دالة الجيب بشكل تكون فيه دالة في f' وتكون بدلالة الزمن **t** (بدلاً عن الزاوية θ)، كما تظهر فيها زاوية الإزاحة وقوة الموجة واللتين يمكن تقديرهما بطريقة المربعات الصغرى حسب طبيعة السلسلة الزمنية.

بما أن السلسلة الزمنية بها وحدات زمنية متقطعة وليست هناك زوايا فيجب تحويل وحدة القياس لوحدة زمنية بدلاً عن زاوية. ولأن السلسلة الزمنية تأخذ قيمها في الأزمنة **n** , **n**- ١ , ... , **t** , ... , ٢ , ١ ,٠ وبما أن مدي قيم الزاوية من 0 الى 2π فإن الزوايا المقابلة لهذه الوحدات الزمنية هي بالترتيب

$$0,\left(\frac{1}{n}\right)2\pi,\left(\frac{2}{n}\right)2\pi,\ldots,\left(\frac{t}{n}\right)2\pi,\ldots,\left(\frac{n-1}{n}\right)2\pi,\ldots,2\pi$$

أي أننا نأخذ قيماً للزوايا بين ٠ و 2π تتزايد بمقدار ثابت $\frac{1}{n}$. وبالتالي يمكن كتابة دالة الجيب بالشكل :

$$Y_t = \sin\left[\left(\frac{t}{n}\right)2\pi\right]$$

حيث المقدار داخل القوس المربع يمثل الزاوية في الزمن t.

لكن هذا يعنى أن قيمة قوة الموجة ستكون فقط بين١+ (عندما تكون الزاوية داخل القوس المربع $\frac{\pi}{2}$) و ١- (عندما تكون $\frac{3\pi}{2}$). و لإفساح المجال لقوة الموجة أن تكون أي مقدار تحدده طبيعة السلسلة نضع

$$Y_t = A\sin\left[\left(\frac{t}{n}\right)2\pi\right]$$

هذا يجعل قوة الموجة تتراوح بين A– وA+.

من ناحية أخري ، من خصائص موجة الجيب أننا إذا ضربنا 2π في العدد الصحيح الموجب f' فإن طولها يتقلص إلي $\frac{n}{f'}$ ما يعني أنها تكمل f' دورة كاملة (أو تكرر f' مرة) خلال الn مشاهدة. نقول في هذه الحالة أن تكرار الموجة f' . وعليه لتكون موجة الجيب ذات تكرار f' كما نرغب نضرب 2π في f' لنحصل على :

$$Y_t = A\sin\left[\left(\frac{f't}{n}\right)2\pi\right]$$

أخيراً ، قد لا تبدأ الموجة الخاصة بالسلسلة من الصفر. لهذا نضيف زاوية زحزحة بمقدار ϕ يتحدد حسب طبيعة السلسلة. تصبح دالة الجيب الآن بالصورة : (٣,١)...

$$Y_t = A\sin\left[\left(\frac{f't}{n}\right)2\pi + \phi\right]$$

هذا يجعل الموجة لا تبدأ من الصفر حتى إذا كانت $t = 0$. مثلاً إذا كانت $\phi = \frac{\pi}{2}$ و t صفر فإن قوة الموجة تصبح A+ وتكون على المحور الراسي.

لدينا الآن دالة تمثل موجة جيب بقوة A ، تكرار f' وزاوية زحزحة ϕ. غير أنه من غير المتوقع أن يتطابق نمط السلسلة الزمنية مع منحني هذه الدالة تماماً ، وذلك بسبب التغيرات العشوائية والطارئة التي يمكن أن تؤثر على قيم السلسلة لهذا سنفترض أن السلسلة الزمنية يمكن تمثيلها بالنموذج التالي :

$$y_t = A\ \sin\left[\left(\frac{f't}{n}\right)2\pi + \phi\right] + e_t \qquad ...(٣,٢)$$

حيث e_t متغير عشوائي بمتوسط صفر وتباين σ^2 وحيث y_t متغير يمثل المتغير Y كانحراف من متوسطه . المعادلة (٣,٢) هي معادلة انحدار غير خطية ويصعب حلها مباشرة بسبب وجود علامة + داخل القوس المربع . ونتذكر أن المجاهيل المراد تقديرها هي A و ϕ .

لتطبيق طريقه المربعات الصغرى على (٣,٢) نستفيد من النتيجة في حساب المثلثات التي تقول :

$$\sin(U + V) = \sin U \cos V + \cos U \sin V$$

وبوضع $V = \phi, U = \left(\frac{f't}{n}\right)2\pi$

تصبح المعادلة :

$$y_t = A\left[(\sin\left(\frac{f't}{n}\right)2\pi)\ (\cos\phi) + (\cos\left(\frac{f't}{n}\right)2\pi)\ (\sin\phi)\right] + e_t$$

$$= b_1\ \sin\left[\left(\frac{f't}{n}\right)2\pi\right] + b_2\cos\left[\left(\frac{f't}{n}\right)2\pi\right] + e_t$$

$$= b_1x_1 + b_2x_2 + e_t \qquad \text{...(٣,٣)}$$

حيث : $b_2 = A\sin\phi \quad , \ b_1 = A\cos\phi$

$$x_1 = \sin\left[\left(\frac{f't}{n}\right)2\pi\right] \quad , x_2 = \cos\left[\left(\frac{f't}{n}\right)2\pi\right]$$

وهي بهذا الشكل تمثل نموذج انحدار عادي يمكن تطبيق طريقة المربعات الصغرى عليـه لإيجاد مقدرات لـ b_1 و b_2 ومن ثم لA و ϕ.

فإذا كانت مقدرات المربعات الصغرى ل b_1 و b_2 هما بالترتيب $\hat{b}_1$ و $\hat{b}_2$

فبما أن $b_1^2 + b_2^2 = A^2\left(\cos^2\phi + \sin^2\phi\right) = A^2$

فإن مقدر قوة الموجة يكون

$$\hat{A} = \sqrt{\hat{b}_1^2 + \hat{b}_2^2}$$

أما $\hat{\phi}$ فيمكن إيجادها بأخذ المقابل لـ$\cos\hat{\phi}$ المعرفة ب

$\cos\hat{\phi} = \dfrac{\hat{b}_1}{\sqrt{\hat{b}_1^2 + \hat{b}_2^2}}$. أو المقابل $\sin\hat{\phi}$ التي تساوي $\dfrac{\hat{b}_2}{\sqrt{\hat{b}_1^2 + \hat{b}_2^2}}$

٣,٢,٢ توفيق k موجة جيب بتكرارات معروفة

يمكن تعميم النتائج السابقة مباشرة لتشمل توفيق عدد k من دوال الجيب حيث الحد الأعلى لk هو $(n-1)/2$ في حالة n عدد فردي و $(n-1)/2$ في حالة عدد زوجي.

النموذج في هذه الحالة تعميم ل(٣,٣) ويأخذ الشكل :

... (٣,٤)

$$y_t = \sum_{i=1}^{k} \left[b_{1i} \ \sin\left(\frac{f_i' t}{n}\right) 2\pi \ + b_{2i} \ \cos\left(\frac{f_i' t}{n}\right) 2\pi) \ \right] + e_t$$

وبتطبيق طريقة المربعات الصغرى نحصل على المقدرات ($\hat{b}_{1i} \ \hat{b}_{2i} ,$) حيث ل $i = 1, 2, ..., k$ ($b_{1i} \ b_{2i} ,$). وبالتالي المقدرات :

$$\hat{A}_i = \sqrt{\hat{b}_{1i}^2 + \hat{b}_{2i}^2} \qquad i = 1, ..., k$$

كما نحصل على ال $\hat{\phi}_i$ من $\sin \hat{\phi}_i = \dfrac{\hat{b}_{2i}}{\hat{A}_i}$

٣,٣ البيريودوقرام The periodogram

افترضنا حتى الآن أن التكرارات f_i' معروفة. لكن في الواقع لا نعرف عادة ما هي الموجات المؤثرة على السلسلة وبالتالي تكراراتها. في هذه الحالة لا نفترض مسبقاً تكرارات معينة ونترك للبيانات بالسلسلة تحديد التكرارات المؤثرة والتي نحتاج لتوفيقها.

إذا وضعنا $f_i' = i$ و $f_i = \dfrac{i}{n}$ يمكن كتابه (٣,٤) بالشكل المختلف قليلاً

$$y_t = \alpha_0 + \sum_{i=1}^{k} \left[\alpha_i \ s_{it} + \beta_i \, c_{it} \right] + e_t \ \ldots (٣,٥)$$

حيـــــث $c_{it}=\cos 2\pi\ f_i t$ و $s_{it}=\sin 2\pi\ f_i t$ وحيـــــث $\alpha_i\ (i=0,1,\ldots,k)$ و $\beta_i\ (i=1,\ldots,k)$ معاملات فـورير. ويطلـق علـى عمليـة تطبيق طريقه المربعات الصغرى لتقدير هذه المعاملات أسماء مختلفة (رغم أن كل منها يحمـل معنـاً معينـاً) منهـا تحليـل فـورير **Fourier analysis** والتحليـل التـوافقي **harmonic analysis** والتحليـــل الطيفـــى **spectral analysis** وتحليـــل البيريودوقرام.

ويمكن إثبات أنه إذا كانت **n** عدد فـردي فـإن مقـدرات المربعـات الصـغرى لـ α_0، $\alpha_i\ (i=1,\ldots,k)$ و $\beta_i\ (i=1,\ldots,k)$ هى بالترتيب:

$$a_0=\frac{1}{n}\sum_{t=1}^{n}Y_t$$

$$a_i=\frac{2}{n}\sum_{t=1}^{n}Y_t s_{it}\qquad i=1,\ldots,k$$

$$b_i=\frac{2}{n}\sum_{t=1}^{n}Y_t c_{it}\qquad i=1,\ldots,k$$

أما إذا كانت **n** عدد زوجي ووضعنا $n=2k$ فإن

$$b_k=\frac{1}{n}\sum_{t=1}^{n}(-1)^t Y_t$$

و $a_k=0$

وتبقي بقية المقدرات كما في حالة **n** عدد فردي.

ويتكون البيريودوقرام من القيم

$$I(f_i)=\frac{n}{2}\left(a_i^2+b_i^2\right)\qquad i=1,\ldots,k$$

في حالة **n** عدد فردي . وفي حالة **n** عدد زوجـي تكـون قيمـة البيريـودوقرام للموجة ذات التكرار **k** :

$$I(f_k)=I(0.5)=nb_k^2$$

وقد تم وضع $f_k = 0.5$ لأن أكبر تكرار نسبي هو ٠,٥ . ذلك لأن أقل عدد من الوحدات الزمنية تحتاجها الموجة لتكمل موجة كاملة هو ٢ ، فإذا كان عدد الوحدات مثلاً ٤٨ فإن أكبر عدد ممكن من الموجات الكاملة سيكون ٢٤ وبالتالي أكبر تكرار نسبي يكون $0.5 = \frac{24}{48}$.

وفي البيريودوقرام تمثل $I\left(f_i\right)$ قوة أو كثافة الموجة ذات التكرار (النسبي) f_i وبالتالي تكون الموجة ذات القوة $I\left(f\right)$ الأكبر هى الأكثر تأثيراً على السلسلة الزمنية. وليس ذلك بمستغرب إذا علمنا أن $I\left(f_i\right)$ تمثل في الواقع مجموع المربعات الخاص بالموجة ذات التكرار f_i إذا أجرينا تحليل تباين جزأنا فيه مجموع مربعات انحرافات قيم السلسلة Y عن وسطها الحسابي $\overline{Y}$. ذلك أنه يمكن إثبات أن مجموع المربعات $\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)^2$ يساوي $\sum_{i=1}^{k} I\left(f_i\right)$. أي أنه يمكن تجزئته إلى $k = \frac{(n-1)}{2}$ مكوّن (في حالة n عدد فردي) كل منها يمثل مجموع المربعات الخاص بزوج $\left(a_i , b_i\right)$ أي $I\left(f_i\right)$ بدرجات حرية ٢. وفي حالة n عدد زوجي يكون هناك $\frac{(n-2)}{2}$ مجموع مربعات كل منها بدرجات حريه ٢ إضافة لدرجة حرية واحدة مرتبطة ب b_k. ونرى من ذلك أن الموجه التى تكون قيمة البيريودوقرام المقابلة لها كبيرة هي التي يكون تأثيرها على التغيرات في السلسلة الزمنية كبيراً.

من ناحية أخري إذا كانت السلسلة الزمنية لا تحوى أي موجات جيب بحيث تتكون كل قيمة Y_t من متوسط عام α_0 وخطأ عشوائي e_t فقط وكانت ال e_t مستقلة وكل منها يتبع التوزيع الطبيعي بوسط حسابي صفر وتباين σ^2 فإن $I\left(f_i\right)$ ستتبع التوزيع $\sigma^2 \chi(2)$ حيث $\chi^2(2)$ توزيع χ^2 بدرجات حرية ٢ وحيث ال $I\left(f_i\right)$ مستقلة.

أما إذا كانت هناك موجة جيب بتكرار f_i وقوة موجة **A** وزاوية زحزحة ϕ فإن قيمة Y_t تأخذ الشكل :

$$Y_t = \alpha_0 + A\left[\cos(2\pi\ f_i\,t)\sin\phi + \sin(2\pi\ f_i\,t)\cos\phi\right] + e_t$$

وفي هذه الحالة يكون توقع $I\left(f_i\right)$ مساويا ل $2\sigma^2 + \frac{n}{2}\left(\alpha^2 + \beta^2\right)$ وليست $2\sigma^2$ كما في حالة السلسلة العشوائية (نذكر أن $\alpha^2 + \beta^2 = A^2$).

مما يعني أن البيرودقرام يتضخم في حالة وجود مكوّن جيب.

٣,٤ طيف العينة **The sample spectrum**

في البيرودوقرام افترضنا أن التكرارات النسبية تأخذ القيم $\frac{1}{n},\frac{2}{n},\frac{3}{n},...$ أي هي مضاعفات التكرار الأساسي **fundamental frequency** $\frac{1}{n}$. إذا تخلينا عن هذا الافتراض وسمحنا ل f بأن تكون متغيراً متصلاً يمكن أن يأخذ أي قيمة في المجال $0-0.5$ فيمكن أن نحصل على الصيغة المعدلة للبيريودوقرام:

$$I\left(f\right) = \frac{n}{2}\left(a_f^2 + b_f^2\right) \qquad\qquad 0 \le f \le 0.5$$

تسمى $I\left(f\right)$ في هذه الحالة طيف العينة. ويستخدم طيف العينة أيضاً لمعرفة الموجات المؤثرة في السلسلة الزمنية وقياس قوتها. وهو الخيار المناسب إذا كنا لا نعلم أن f يمكن أن تأخذ فقط إحدى القيم $\frac{1}{n},\frac{2}{n},\frac{3}{n},...$ بل يمكن أن تأخذ أى قيمة في المدى ٠ إلي ٠,٥ .

٣,٥ الطيف **The spectrum** ودالة كثافة الطيف **Spectral density function**

في السلاسل الزمنية التي تكون خاضعة لموجات جيب ذات تكرارات محددة يساعد كل من البيريودوقرام وطيف العينة في إبراز الموجات المؤثرة. ولكن هناك سلاسل زمنية تتغير فيها تكرارات وقوة وزحزحه الموجات بشكل عشوائي. في مثل

هذه السلاسل يُظهر كل من البيريودوقرام وطيف العينة تقلبات كبيرة بحيث لا يمكن إعطاء قيمتها معنى.

أفرض الآن أن لدينا سلسلة زمنية بحجم **n** وأن هذه السلسلة يمكن النظر إليها كتحقيق (في الواقع) لعملية تتبع التوزيع الطبيعي ولا تتغير مع الزمن. إذا

أجرينا عدداً من التحقيقات لهذه العملية بحيث يتكون كل تحقيق من **n** مشاهدة فيمكن حساب a_f, b_f و $I(f)$ لكل تحقيق (أو سلسلة زمنية تمت مشاهدتها).

وإذا رمزنا لمتوسط قيم $I(f)$ ب $E(I(f))$ فإن نهاية هذا المتوسط عندما تؤول **n** لما لا نهاية تسمي طيف القوة **Power Spectrum** ويرمز له ب$P(f)$ أي

$$P(f) = \lim_{n\to\infty} E(I(f))$$

ويشار لطيف القوة عادة بالطيف **spectrum** اختصاراً .

إذا قسمنا الطيف $P(f)$ على تباين فيم السلسلة σ_y^2 نحصل على ما يسمي بدالة كثافة الطيف ونرمز لها ب $K(f)$:

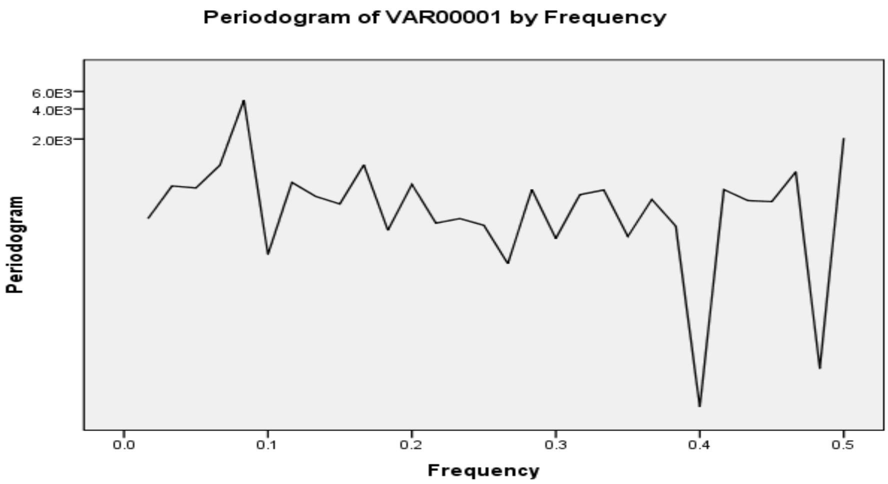
$$K(f) = P(f)/\sigma_y^2$$

وتتميز $K(f)$ بخصائص دالة الكثافة الاحتمالية

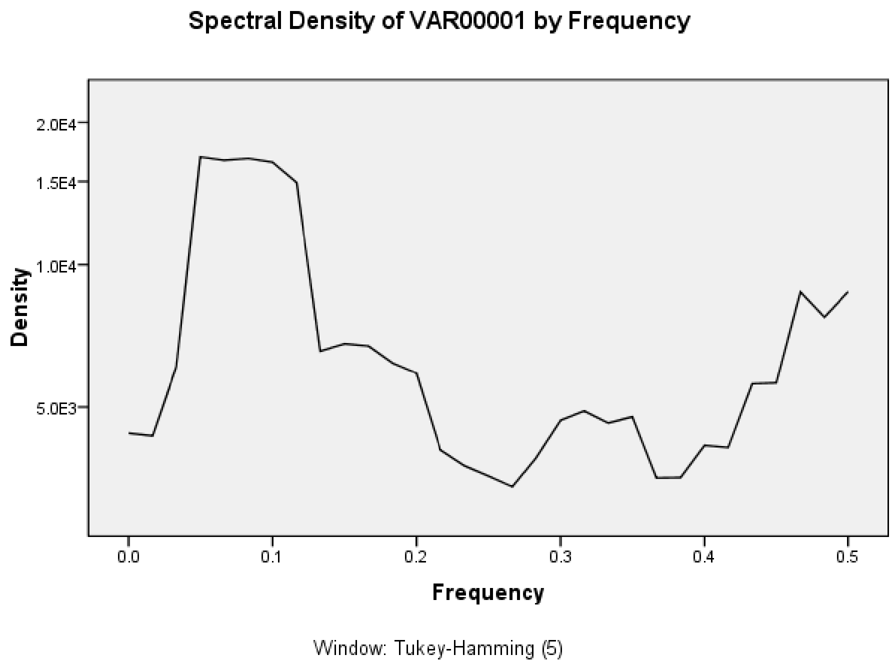
$$K(f) > 0, \qquad \int_0^{0.5} K(f)df = 1$$

مثال (١ ,٣)

يوضح شكل (١ ,٣) و(٢ ,٣) البيريودوقرام وكثافة الطيف بالترتيب لبيانات سلسلة الزلاجات المائية. ونلاحظ من الشكل (١ ,٣) أن اكبر موجه هى ذات التكرار ٠٨ , ٠ تقريباً مما يعنى أنها لموجه ذات طول ١٢ تقريباً(مقلوب ٠٨ , ٠).


Periodogram of VAR00001 by Frequency
6.0E3
4.0E3
2.0E3
Periodogram
0.0
0.1
0.2
0.3
0.4
0.5
Frequency

شكل (١ , ٣)


Spectral Density of VAR00001 by Frequency
2.0E4
1.5E4
1.0E4
5.0E3
Density
0.0
0.1
0.2
0.3
0.4
0.5
Frequency
Window: Tukey-Hamming (5)

٣,٦ دالة التغاير الذاتي The autocovariance function

من المفاهيم المرتبطة ارتباطاً وثيقاً بالتحليل الطيفي والتي لها أهمية كبيرة في بعض نماذج التنبؤ التي ستتعرض لها لاحقاً مفهوم التغاير الذاتي ومفهوم الارتباط الذاتي. والتعاريف التالية لهذه المصطلحات تفترض أن العملية التصادفيه التي تولدت عنها السلسلة الزمنية ذات متوسط وتباين ثابت لا يتغير مع الزمن. ويعرّف التغاير الذاتي للمجتمع (تحديداً للعملية التصادفيه المفترض أنها ولدت السلسلة الزمنية المشاهدة) بإبطاء k (lag k):

$$\gamma_k = \operatorname{cov}\left[Y_t, Y_{t+k}\right]$$

$$= E\left[\left(Y_t - \mu\right)\left(Y_{t+k} - \mu\right)\right] \qquad \dots (٣,٦)$$

حيث μ متوسط المجتمع و cov تعني تغاير. بمعني آخر التغاير الذاتي بإبطاء k هو التغاير بين القيم التي تبعد عن بعض k وحده زمنية وواضح أن γ_0 يُمثل التباين.وإذا نظرنا للتغاير الذاتي للمجتمع كدالة في الإبطاء k تكون لدينا دالة التغاير الذاتي للمجتمع **population autocovariance function.**

وبما أننا في الواقع نشاهد سلسلة زمنية محدودة بحجم n مثلاً فاننا نحتاج لتقدير γ_k منها. وهناك عدة مقدرات للتغاير بإبطاء k أكثرها استخداماً المقدر

$$C_k = \frac{1}{n}\sum_{t=1}^{n-k}\left(Y_t - \bar{Y}\right)\left(Y_{t+k} - \bar{Y}\right) \qquad \dots (٣,٧)$$

وكدالة في k تعطى قيم C_k دالة تغاير العينة **sample autocovariance function**

٣,٧ الإرتباط الذاتي The autocorrelation

معامل الارتباط الذاتي عندما يكون المتوسط ثابت والتباين غير ثابت يعرف :

$$\rho_k = \frac{E\left[\left(Y_t - \mu\right)\left(Y_{t+k} - \mu\right)\right]}{\sqrt{E\left(Y_t - \mu\right)^2 \; E\left(Y_{t+k} - \mu\right)^2}}$$

وإذا كان التباين أيضاً ثابتاً فإن $E\left(Y_t - \mu\right)^2 = E\left(Y_{t+k} - \mu\right)^2$

وبالتالي يكون الارتباط الذاتي بإبطاء k:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \qquad \ldots (٣,٨)$$

وإذا نظرنا ل ρ_k كدالة في k نحصل على دالة الارتباط الذاتي autocorrelation function (ACF) ويقدر معامل الارتباط الذاتي من العينة ب r_k حيث

$$r_k = \frac{\sum_{t=1}^{n-k}\left(Y_t - \bar{Y}\right)\left(Y_{t+k} - \bar{Y}\right)}{\sum_{t=1}^{n}\left(Y_t - \bar{Y}\right)^2} \qquad \ldots (٣,٩)$$

بافتراض ثبات المتوسط والتباين

٣,٦ العلاقة بين طيف العينة ومقدر التغاير الذاتي

يرتبط طيف العينة $I(f)$ بمقدر التغاير الذاتي C_k بعلاقة هامة نوردها في النتيجة التالية.

نتيجة (٣,١)

$$I(f) = 2\left\{C_0 + 2\sum_{k=1}^{n-1} C_k \cos 2\pi f k\right\} \qquad 0 \le f \le 0.5$$

الإثبات

$$I(f) = \frac{n}{2}\left(a_f^2 + b_f^2\right) = \frac{n}{2}\left(b_f - ia_f\right)\left(b_f + ia_f\right)$$

حيث $i = \sqrt{-1}$ **وباستخدام**

$$\left.\begin{aligned} a_i &= \frac{2}{n}\sum_{t=1}^{n} Y_t s_{it} \\ b_i &= \frac{2}{n}\sum_{t=1}^{n} Y_t c_{it} \end{aligned}\right\} i = 1,\ldots,k$$

ووضع المشاهدات في شكل انحراف عن المتوسط* :

$$\left(b_f - ia_f\right) = \frac{2}{n}\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)\left(c_{it} - is_{it}\right)$$

$$= \frac{2}{n}\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)\left(\cos 2\pi ft - i\sin 2\pi ft\right)$$

$$= \frac{2}{n}\sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)e^{-i2\pi ft}$$

بالاستعانة بمعادلة إيولر $e^{-iz} = \cos z - i\sin z$.

وبالتالي:

$$I(f) = \sum_{t=1}^{n}\left(Y_t - \overline{Y}\right)e^{-i2\pi ft} \times \frac{2}{n}\sum_{t'=1}^{n}\left(Y_{t'} - \overline{Y}\right)e^{i2\pi ft'}$$

$$= \frac{2}{n}\sum_{t=1}^{n}\sum_{t'=1}^{n}\left(Y_t - \overline{Y}\right)\left(Y_{t'} - \overline{Y}\right)e^{-i2\pi f(t-t')}$$

إذا وضعنا $k = t - t'$ ينتج بما أن $t = t' + k$ وأقصي قيمة لـt هى n أن t' في المجموع الثاني لن تكون أكبر من $n-k$، وبالتالي

$$I(f) = \frac{2}{n}\sum_{t=1}^{n}\sum_{t'=1}^{n-k}\left(Y_t - \overline{Y}\right)\left(Y_{t'} - \overline{Y}\right)e^{-i2\pi fk}$$

$$= 2\sum_{t=1}^{n}\left\{\frac{1}{n}\sum_{t'=1}^{n-k}\left(Y_{t'+k} - \overline{Y}\right)\left(Y_{t'} - \overline{Y}\right)\right\} e^{-i2\pi fk}$$

ولكن من (٣,٧) المقدار داخل القوس المتعرج هو C_k. إذن

$$I(f) = 2\sum_{t=1}^{n} C_k e^{-i2\pi fk}$$

* هذا يكافئ نقل α_0 من الطرف الأيمن من النموذج (٣.٥) ولا يؤثر في المقدرات b,a.

من التحويله $k=t-t'$ نلاحظ أنه عندما تكون $t=1$ تصبح $k=1-t'$ وتأخذ k قيمتها الصغرى عندما تأخذ t' قيمتها الكبرى **n**. هذا يعني أن أصغر قيمة لـ**k** هي $k=1-n=-(n-1)$

كذلك عندما تكون $t=n$ تصبح $k=n-t'$ وتأخذ **k** قيمتها الكبرى عندما تأخذ t' قيمتها الصغرى وهي ١. هذا يعني أن أكبر قيمة لـ**k** هي $n-1$. وبالتالي

$$I(f)=2\sum_{k=-(n-1)}^{n-1} C_k e^{-i2\pi fk}$$

$$=2\sum_{k=-(n-1)}^{n-1} C_k\left[\cos(2\pi fk)-i\sin(2\pi fk)\right]\quad 0\le f\le 0.5$$

باستخدام معادلة أيولر مرة أخرى . لكن دالة الجيب دالة فردية بمعني أن $f(-x)=-f(x)$ وبالتالي عند جمعها في الفترة المتماثلة $[-(n-1),(n-1)]$ يكون المجموع صفراً. هذا يؤدي لاختفاء الحد الثاني داخل القوس ونصل للنتيجة:

$$I(f)=2\sum_{k=-(n-1)}^{n-1} C_k\cos(2\pi fk)$$

لكن عند $k=0$ يكون $\cos(2\pi fk)=\cos 0=1$ وبالتالي يمكن فصل الحالة $k=0$ من المجموع

$$I(f)=2\left[C_0+\sum_{\substack{k=-(n-1)\\ k\neq 0}}^{n-1} C_k\cos(2\pi fk)\right]$$

كذلك لتماثل الفترة $[-(n-1),n-1]$ ولأن $\cos(-\theta)=\cos(\theta)$:

$$\sum_{\substack{k=-(n-1)\\ k\neq 0}}^{n-1} C_k\cos(2\pi fk)=2\sum_{k=1}^{n-1} C_k\cos(2\pi fk)$$

لنصل أخيراً للنتيجة المراد إثباتها :

$$I(f)=2\left[C_0+2\sum_{k=1}^{n-1}C_k\cos 2\pi fk\right]\qquad 0\le f\le 0.5$$

وقد وجد أن متوسط مقدر معامل التغاير الذاتي يؤول لمعامل التغاير الذاتي في المجتمع عندما $n\to\infty$ أي

$$\lim_{n\to\infty}E(C_k)=\gamma_k$$

من ذلك نستنتج التالي :

$$P(f)=\lim_{n\to\infty}E(I(f))=2\left[\lim_{n\to\infty}E(C_0)+2\sum_{k-1}^{\infty}\lim_{n\to\infty}E(C_k)\cos(2\pi fk)\right]$$

$$=2\left[\gamma_0+2\sum_{k=1}^{\infty}\gamma_k\cos(2\pi fk)\right]$$

وإذا قسمنا على التباين γ_0 نحصل على دالة كثافة الطيف

$$K(f)=2\left[1\ +2\sum_{k=1}^{\infty}\gamma_k\rho_k\cos(2\pi fk)\right]\qquad 0\le f\le 0.5$$

كدالة في الارتباط الذاتي للمجتمع.

العلاقة بين دالة كثافة الطيف ودالة الارتباط الذاتي تعني أن كل منهما يمكن الحصول عليه كتحويله من الآخر مما يعني أيضاً أنهما متكافئتان رياضياً. ولكن هذا لا يعني أنه يمكن الاستغناء عن أحداهما لأن كل منهما يسلط الضوء على جانب مختلف من السلسلة الزمنية. فدالة كثافة الطيف تلقي الضوء على الموجات المؤثرة والطاغية في السلسلة وتكراراتها بينما توضح دالة الارتباط الذاتي ما إذا كانت القيم المتتالية في السلسلة ترتبط بارتباط موجب (ينعكس في شكل تمهيد نسبي بالسلسلة) أم ارتباط سالب (تظهر فيه السلسلة بشكل تتبادل فيه التغيرات الموجبة والسالبة الظهور). وكما قال بوكس وجنكينز (١٩٧٦) "أنهما معاً تساعدان في جعل السلسلة الزمنية تتحدث

عن نفسها ويلعبان بالتالي في تحليل السلاسل الزمنية الدور الذي يلعبه المدرج التكراري في تحليل توزيع البيانات بإشارته للتوزيع النظري الذي يمكن أن يكون مناسباً لتمثيلها.

مثال (٢ , ٣) :

افترض العملية التصادفيه البسيطة

$$Y_t = 5 + e_t + e_{t-1} + e_{t-2}$$

حيث e_t، e_{t-1} و e_{t-2} متغيرات عشوائية غير مرتبطة يتبع كل منها التوزيع الطبيعي بمتوسط ٠ وتباين ١. يمكن تمثيل هذه العملية من خلال دالة التغاير الذاتي ، أو دالة الارتباط الذاتي أو دالة كثافة الطيف كما يلي :

نلاحظ أولاً أن : $E(e_i) = 0$ لكل i وبالتالي $\sigma^2 = V(e_i) = E(e_i^2) = 1$.

كذلك وبسبب عدم ارتباط ال $e's$ فإنه ل $i \neq i'$:

$$\operatorname{cov}(e_i, e_{i'}) = E(e_i\, e_{i'}) = 0$$

أيضاً $E(Y_t) = 5$

دالة التغاير الذاتي :

عند $k = 0$ (التباين) :

$$\gamma_0 = E[Y_t - 5]^2 = E[5 + e_t + e_{t-1} + e_{t-2} - 5]^2$$
$$= E(e_t^2) + E(e_{t-1}^2) + E(e_{t-2}^2) + 2E(e_t e_{t-1} + e_t e_{t-2} + e_{t-1} e_{t-2})$$
$$= 1 + 1 + 1 + 0 = 3$$

عند $k = 1$:

$$\gamma_1 = E[Y_t - 5][Y_{t+1} - 5] = E(e_t + e_{t-1} + e_{t-2})(e_{t+1} + e_t + e_{t-1})$$
$$= E(e_t^2) + E(e_{t-1}^2) + E(2e_t e_{t-1} + e_t e_{t+1} + e_{t-1} e_{t+1} + e_{t-2} e_{t+1} + e_{t-2} e_t + e_{t-2} e_{t-1})$$
$$= 1 + 1 + 0 = 2$$

عند $k = 2$:

$$\gamma_2 = E[Y_t - 5][Y_{t+2} - 5] = E(e_t + e_{t-1} + e_{t-2})(e_{t+2} + e_{t+1} + e_t)$$
$$= E(e_t^2) = 1$$

ل $k > 2$ تكون مؤشرات ال $e's$ في القـوس الثـاني جميعهـا مختلفـة عـن تلـك الـتي بالقوس الأول وبالتالي يكون التغاير الذاتي ٠. دالة التغاير الذاتي إذن

$$\gamma_k = \begin{cases} 3 & k = 0 \\ 2 & k = 1 \\ 1 & k = 2 \\ 0 & k > 2 \end{cases}$$

دالة الارتباط الذاتي يمكن الحصول عليها من هذه الدالة بالقسمة على γ_0:

$$\rho_k = \begin{cases} \frac{2}{3} & k = 1 \\ \frac{1}{3} & k = 2 \\ 0 & k > 2 \end{cases}$$

دالة كثافة الطيف :

بتعويض قيم ρ_k في

$$K(f) = 2\left\{1 + 2\sum_{k=1}^{\infty} \rho_k \cos 2\pi fk\right\}$$
$$= 2\left\{1 + \frac{4}{3}\cos 2\pi f + \frac{2}{3}\cos 4\pi f\right\}$$

مـع ملاحظـة أن حـدود المجمـوع المقابلـة لأى $k > 2$ تتلاشـى لأن $\rho_k = 0$ ل $k > 2$

الباب الرابع

طرق التمهيد

Smoothing methods

٤,١ مقدمة

في الباب الثاني والثالث تحدثنا عن طرق تساعد في فهم طبيعة السلسلة الزمنية من خلال عزل وقياس (ما يمكن قياسه من) التغيرات المختلفة التي تؤثر فيها ، أو اكتشاف أي موجات تحتويها وتحديد أطوالها وتكرارها. ويمثل ذلك أحد الأهداف الأساسية لتحليل السلسلة الزمنية. هدف آخر لا يقل أهمية وراء تحليل السلسلة الزمنية هو الاستفادة من القيم التاريخية بها للتنبؤ بالقيم المستقبلية. وفي هذا الباب والبابين التاليين له سنتعرف على بعض الطرق التي تستخدم في التنبؤ من السلسلة الزمنية. والطرق التي سيتم تناولها هي طرق التمهيد والطرق المستندة إلى مجموعة النماذج التي تدخل فيما يعرف بمنهجية بوكس – جنكينز. وتقوم طرق التمهيد بصفة عامة باستخدام أنواع مختلفة من المتوسطات بهدف تقليص الفوارق بين القيم الكبيرة والقيم الصغيرة في السلسلة الزمنية للوصول لسلسلة جديدة تكون قيمها أقرب لبعضها – أي أكثر تمهيداً – من قيم السلسلة الأصلية. والفكرة الأساسية وراء ذلك ، هي أن التمهيد بقضائه على نسبة كبيرة من التغيرات قصيرة الأمد (كالتغيرات العشوائية والموسمية) يتيح الفرصة لإبراز الاتجاه العام للسلسلة الزمنية مما يمكن من التنبؤ بقيمها المستقبلية في ضوئه. ورغم أن طرق التمهيد لا تستند إلى نظرية إحصائية (مثلها في ذلك مثل طرق التجزئة) وتعتمد أساساً على الحدس والتجربة والمنطق ، إلا أنها أثبتت نجاحاً خاصة في التنبؤ قصير الأمد. وتتميز طرق التمهيد عموماً بأنها تكيفيه **adaptive** بمعنى أن التنبؤ يعدل مع ظهور كل قيمة جديدة مما يجعله مبنياً على صيغة تتطور باستمرار بدلاً من الاعتماد على صيغة أو معادلة ثابتة لا تستوعب ما يتوفر من معلومات جديدة عن السلسلة مع مرور الزمن أولا تستوعبها بدرجة كافية.

٤,٢ طريقة المتوسط The average method

أفرض أن لدينا قيم لسلسلة زمنية حتى الزمن t :

$$Y_1, Y_2, \ldots, Y_t$$

ونرغب في التنبؤ بالقيمة في الزمن $t+1$ ، والذي نرمز له ب F_{t+1}.

إذا لم تكن السلسلة تحوى اتجاهاً عاماً أو تغيرات موسمية ، وتبدو قيمها وكأنها تتشتت عشوائياً حول قيمة وسطى ثابتة ، فإنه يبدو منطقياً استخدام كتنبؤ بالقيمة في الزمن $t+1$ ، مقدر لهذه القيمة الوسطى . وتستخدم طريقة المتوسط متوسط قيم السلسلة حتى الزمن t كتنبؤ أي:

$$F_{t+1} = \frac{1}{t}\sum_{i=1}^{t} Y_i$$

فإذا توفرت بعد ذلك – في الزمن $t+1$ – القيمة الفعلية Y_{t+1} فإن التنبؤ بالقيمة في الزمن $t+2$ يكون :

$$F_{t+1} = \frac{1}{t+1}\sum_{i=1}^{t+1} Y_i$$

وهكذا. وواضح أن هذه الطريقة ليست مناسبة إذا كانت السلسلة الزمنية خاضعة لتغيرات مع الزمن لأنها تعطي نفس الوزن (مثلاً $\frac{1}{t+1}$) للقيم حديثها وقديمها.

٤,٣ طريقة المتوسط المتحرك Moving average method

كمحاولة للتخلص من تأثير القيم القديمة على المتوسط ، ووضع احتمال وجود اتجاه عام في السلسلة في الاعتبار ، يتم أحياناً استخدام طريقة المتوسط المتحرك .

وفي أبسط صورها وهي المتوسط المتحرك المفرد **single moving average** والتي نقوم فيها بحساب متوسطات مجموعات متتالية من قيم السلسلة ، القيم في كل مجموعة عبارة عن القيم في المجموعة السابقة لها بعد حذف أقدم قيمة وإضافة القيمة التي تليها. ويستخدم المتوسط المتحرك في الزمن **t** كتنبؤ بالزمن $t+1$.

فإذا بدأنا مثلاً بالقيم: $Y_1, Y_2, \ldots, Y_t$ فإن التنبؤ بالقيمة في الزمن $t+1$ يكون:

$$F_{t+1} = \frac{1}{t} \sum_{i=1}^{t} Y_i$$

ويتطابق في هذه الحالة مع التنبؤ بطريقة المتوسط . وبمجرد ظهور القيمة الفعلية في الزمن $t+1$ ، Y_{t+1} يكون التنبؤ في الزمن $t+2$:

$$F_{t+2} = \frac{1}{t} \sum_{i=2}^{t+1} Y_i$$

لاحظ أن المجموع في F_{t+2} حذفت منه القيمة الأقدم في F_{t+1} وهى Y_1 وأضيفت القيمة التي تلي آخر قيمة فيه Y_t وهي Y_{t+1}. كذلك

$$F_{t+3} = \frac{1}{t} \sum_{i=3}^{t+2} Y_i$$

وهكذا نتخلص من القيم القديمة الواحدة تلو الأخرى ، ويجدد بذلك التنبؤ ونحن ننتقل مع الزمن . ولكن كيف يتم التجديد أو التعديل ؟

لرؤية ذلك نضع :

$$F_{t+2} = \frac{1}{t} \sum_{i=2}^{t+1} Y_i$$

$$= \frac{1}{t}\left(\sum_{i=2}^{t+1} Y_i + Y_1 - Y_1\right)$$

$$= \frac{1}{t}\left(\sum_{i=1}^{t} Y_i - Y_1 + Y_{t+1}\right)$$

$$= F_{t+1} + \frac{1}{t}\left(Y_{t+1} - Y_1\right) \qquad ...(٤ , ١)$$

أي أن التعديل الذي يتم في كل تنبؤ هو إضافة الفرق بين القيمة التي أضيفت والقيمة التي حذفت مضروباً في $\frac{1}{t}$.

ويمكن أن نختار أي رتبة مناسبة للمتوسط المتحرك . فإذا اخترنا مثلاً متوسط متحرك برتبة **n** فإن التنبؤ في الزمن $t+1$ يكون من (١ , ٤) :

$$F_{t+1} = F_t + \frac{1}{n}\left(Y_t - Y_{t-n}\right) \qquad \ldots (٤ , ٢)$$

طريقة المتوسط المتحرك المفرد ليست هي الطريقة الوحيدة المبنية على فكرة المتوسطات المتحركة. وفي الواقع هناك عدة أشكال للمتوسطات المتحركة مقترحة في الأدبيات. بصفة خاصة ، إذا كان تطبيق المتوسط المفرد على السلسلة الزمنية يؤدي لنوع من الخطأ المنتظم (اتجاه عام خطي يتزايد بمقدار ثابت) فإن طريقة المتوسطات المتحركة الخطية **Linear moving average** التي تتطلب إيجاد متوسط متحرك مضاعف (متوسط متحرك لسلسلة المتوسط المتحرك المفرد) يمكن أن نستخدم لتحسين التنبؤ.

٤ , ٤ طريقة المتوسطات المتحركة الخطية Linear moving averages

أفرض أن لدينا سلسلة زمنية خاضعة لاتجاه عام تتزايد فيه القيم بمقدار ثابت . أفرض أيضاً أننا نريد أن نستخدم متوسط متحرك (مفرد) بطول فترة ٣. تتطلب طريقة المتوسطات المتحركة الخطية إتباع الخطوات التالية :

١. نحسب متوسط القيم الثلاث الأولي ونضعه أمام القيمة الثالثة. لاحظ أن هذا يختلف عما كنا نفعله في الباب الثاني حين كنا نضع المتوسط المتحرك أمام القيمة التي في الوسط أو عندما نستخدمه كتنبؤ فنضعه أمام القيمة الرابعة.

نرمز للمتوسط المتحرك ، والذي هو متوسط متحرك مفرد ، أمام القيمة رقم **t** (أو الزمن **t**) ب S'_t. لاحظ أن

$$S'_t = \frac{1}{3} \sum_{i=t-3+1}^{t} Y_i$$

٢. نحسب متوسط متحرك مضاعف (أي متوسط متحرك لقيم المتوسط المتحرك المفرد S'_t) برتبة ٣ أيضاً، ونضع كل متوسط متحرك مضاعف أمام آخر متوسط متحرك مفرد دخل في حسابه.

نرمز للمتوسط المضاعف مقابل الزمن **t** ب S''_t حيث

$$S_t'' = \frac{1}{3} \sum_{i=t-3+1}^{t} S_i'$$

٣. التنبؤ لـm فترة زمنية للأمام إذا كنا نقف في الزمن t يحسب من

$$F_{t+m} = a_t + b_t m$$

حيث:

$$a_t = S_t' + \left(S_t' - S_t''\right)$$

و $$b_t = \frac{2}{3-1}\left(S_t' - S_t''\right)$$

يمثل الاتجاه العام.

وفي الحالة العامة عند استخدام متوسطات متحركة برتبة r تكون هذه المعادلات بالترتيب

$$F_{t+m} = a_t + b_t m$$

$$a_t = S_t' + \left(S_t' - S_t''\right)$$

$$b_t = \frac{2}{r-1}\left(S_t' - S_t''\right)$$

حيث

$$S_t'' = \frac{1}{r} \sum_{i=t-r+1}^{t} S_i' \quad \text{و} \quad S_t' = \frac{1}{r} \sum_{i=t-r+1}^{t} Y_i$$

والسبب في وجود العامل $\frac{2}{r-1}$ عند حساب مكوّن الاتجاه العام b_t هو أن المتوسط المتحرك ذو الرتبة r $\left(S_t'\right)$ من المفروض أن يوضع (كما ذكرنا في الباب الثاني) أمام القيمة التي في الوسط أي أمام الفترة الزمنية $\frac{r+1}{2}$ (في حالة أول متوسط متحرك) بينما وضع في الواقع أمام الفترة r . إذن هناك فرق يساوي $\frac{r-1}{2} = r - \frac{r+1}{2}$. نفس الفرق ينطبق على المتوسط المتحرك المضاعف

$\left(S_t''\right)$ وبالتالي فإن الفرق (الاتجاه العام) $S_t' - S_t''$ يمثـل الفـرق ل $\frac{r-1}{2}$ وحـدة زمنية مما يعني أن الاتجاه العام للوحدة الزمنية الواحدة هو $\frac{2}{r-1}\left(S_t' - S''\right)$.

المثال التالي يوضح أنه إذا كانت السلسلة الزمنية خاضعة لاتجاه عام ثابت فإن طريقـة المتوسطات المتحركة الخطية تؤدي للتنبؤ بدون خطأ (وبالتالي تقلل الخطأ إذا كان شبه ثابت).

مثال (١ , ٤)

جدول (١ , ٤) يوضح سلسلة افتراضية يتزايـد فيهـا الاتجـاه العـام بمقـدار ثابـت (٢) والحسابات المطلوبة للتنبؤ بطريقة المتوسطات المتحركة الخطية :

الزمن t	السلسلة Y_t	M(٣) $\left(S_t'\right)$	(3×3) $\left(S_t''\right)$	$-S_t''$	$S_t' + \left(S_t' - S_t''\right)$	$+\frac{2}{3-1}\left(S_t' - S_t''\right)$	$= a_t + b_t$
١	٤						
٢	٦						
٣	٨	٦					
٤	١٠	٨					
٥	١٢	١٠	٨	٢	١٢	٢	
٦	١٤	١٢	١٠	٢	١٤	٢	١٤
٧	١٦	١٤	١٢	٢	١٦	٢	١٦

جدول (١ , ٤)

نجد من الجدول أن التنبؤ بالقيمة في الزمن $t = 6$ إذا كنا في الزمن ٥ :

$$F_6 = S_5' + \left(S_5' - S_5''\right) + \frac{2}{3-1}\left(S_5' - S_5''\right)$$

$$= 10 + 2 + 2 = 14$$

وهي تتطابق مع القيمة الفعلية $Y_6 = 14$ مما يعني أن التنبؤ تم بدون خطأ.

وكذلك للتنبؤ F_7 . أما إذا اكتفينا بالمتوسط المتحرك المفرد (أي استخدمنا مثلاً $5_5''$ لتنبؤ ب Y_6 فإن خطأ منتظم بمقدار ٤ كان سيحدث عن التنبؤ.

٤ , ٥ طرق التمهيد الأسى Exponential smoothing methods

لقد وجدنا أن طريقة المتوسط تعطي أوزاناً متساوية لقيم السلسلة قديمها وحديثها ، وأنها بذلك لا تناسب السلاسل الزمنية التي تخضع لتغيرات مع الزمن. وفي طريقة المتوسطات المتحركة محاولة لمعالجة المشكلة من خلال التخلص من القيم القديمة الواحدة تلو الأخرى إذ تسقط أقدم القيم في المتوسط المتحرك السابق وتستبدل بأحدث قيمة عند حساب المتوسط المتحرك الجديد .

طرق التمهيد الأسى أيضاً تحاول التركيز على القيم الأحدث ولكن عن طريق منح أوزان مختلفة لبيانات السلسلة الزمنية بحيث يتناقص وزن القيمة كلما قدمت. وقد وجد أن هذه الطرق تعطي نتائج جيدة عندما تكون السلسلة الزمنية عرضه لتغيرات بطيئة . في مثل هذه الحالة من الضروري استخدام طريقة للتنبؤ تنطوي علي تجديد وتطوير التنبؤ كلما ظهرت قيمة جديدة ، إذا كان لها أن تأخذ أحدث التغيرات في السلسلة الزمنية في الاعتبار.

وكما هو الحال في الطرق المبنية على المتوسطات المتحركة ، توجد طرق عديدة للتمهيد الأسى. وسنتناول ثلاث من هذه الطرق وهي طرق التمهيد الأسى المفرد ، التمهيد الأسى الثنائي لهولت والتمهيد الأسى الثلاثي لوينترز . وهذه هي طرق التمهيد الأسى الأكثر شهرة.

٤ , ٥ , ١ التمهيد الأسى المفرد Single exponential smoothing

إذا كانت السلسلة الزمنية غير خاضعة لتأثير اتجاه عام أو تأثير موسمي ، فيمكن اعتبار قيمها ناتجة عن متوسط عام مضافاً إليه خطأ عشوائي تختلف قيمته من زمن لآخر. ويبدو منطقياً في هذه الحالة استخدام مقدر مناسب لهذا المتوسط العام عند التنبؤ بقيمة جديدة للسلسلة. وبما أن المتوسط يمكن أن يتغير ببطء فإن التمهيد الأسى يقوم بإعطاء وزن أكبر للقيم الأحدث في السلسلة عند حساب المقدر كما أنه يجدد

ويحدّث المقدر كلما ظهرت قيمة جديدة في السلسلة الزمنية. وتأخذ المعادلة الأساسية للتمهيد الأسى المفرد الشكل :

$$F_{t+1} = \alpha Y_t + (1-\alpha) F_t \qquad ...(٤,٣)$$

حيث F_t التنبؤ بالقيمة في الزمن t إذا كنا في الزمن $t-1$ ، Y_t القيمة الفعلية في السلسلة الزمنية في الزمن t . أما α والتي تتراوح قيمتها بين ٠ و١ فهي ثابت تمهيد **smoothing constant** . وتحدد قيمة α عادة بمحاولة قيم مختلفة ثم اختيار القيمة التي تعطي أقل مجموع (أو متوسط) مربعات خطأ .

بعض البرمجيات تختار قيم ل α بين 0.01 و 0.3 بقفزات 0.01. أي نقوم بتجربة 0.01 ، 0.02، 0.03،... حتى ٠,٣ . لاحظ أنه كلما كانت التغييرات في السلسلة كبيرة كلما احتجنا ل α أكبر. لاحظ كذلك الشبة بين (٤,٣) و (٤,٢).

وتتميز المعادلة (٤,٣) بأننا لا نحتاج للتنبؤ بقيمة جديدة ، سوى الاحتفاظ بآخر قيمة مشاهدة وآخر تنبؤ وتحديد قيمة مناسبة لثابت التمهيد α. ولكن ماذا يفعل التمهيد الأسى تحديداً ؟

إذا كتبنا (٤,٣) بالشكل

$$\begin{aligned} F_{t+1} &= F_t + \alpha(Y_t - F_t) \\ &= F_t + \alpha(\delta_t) \end{aligned}$$

حيث δ_t خطأ التنبؤ في الزمن t ، نتبين أن ما يفعله التمهيد الأسى عند التنبؤ بالقيمة في الزمن $t+1$ هو أن يأخذ التنبؤ السابق F_t ويصححه مستهدياً بخطأ التنبؤ في الزمن t. δ_t ويتم التصحيح في اتجاه معاكس لاتجاه الخطأ. فمثلاً إذا كان التنبؤ في الزمن t أكبر من الواقع فإن الخطأ $(Y_t - F_t)$ (والذي سيكون سالباً في هذه الحالة) سيضاف بعد ضربه في α إلى التنبؤ السابق F_t فيكون التنبؤ الجديد F_{t+1} أقل من F_t . وهذه هو التصحيح. أما إذا كان التنبؤ في الزمن t أقل من الواقع فإن إضافة الخطأ (الذي سيكون موجباً) بعد ضربه في α ل F_t سيقلل من الخطأ في التنبؤ الجديد. هذا يعني أن التمهيد الأسى يتضمن نوعاً من التغذية الاسترجاعية السلبية.

من ناحية أخرى بتكرار استخدام القاعدة في (٣ , ٤) نجد :

$$F_{t+1} = \alpha Y_t + (1-\alpha)(\alpha Y_{t-1} + (1-\alpha)F_{t-1})$$
$$= \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + (1-\alpha)^2 F_{t-1}$$
$$= \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \alpha(1-\alpha)^3 Y_{t-3} + \ldots$$
$$\ldots + \alpha(1-\alpha)^{n-1} Y_{t-(n-1)} + \ldots + (1-\alpha)^n F_{t-(n-1)}$$

وبما أن α كسر تتراوح قيمته بين الصفر والواحد فواضح أن التمهيد الأسى يعطي وزناً أقل للقيم الأقدم . في الواقع فإن تناقص الأوزان يتبع نمطاً أسياً ولهذا التسمية.

ليمكن استخدام المعادلة (٣ , ٤) في التنبؤ نحتاج لنقطة ننطلق منها وتحديداً نحتاج لقيمة ل F_1 التنبؤ بالقيمة في الزمن ١ إذ لا يمكننا استخدام (٣ , ٤) لإيجاد F_1 لعدم وجود قيمة Y في الزمن ٠. أثبتت التجارب أنه في التمهيد الأسى عموماً يمكن استخدام متوسط نصف قيم السلسلة كتقدير ل F_1 وأنه في التمهيد الأسى المفرد يمكن الاكتفاء بأخذ متوسط القيم ال٦ الأولي أي نأخذ :

$$F_1 = \frac{1}{6}\sum_{t=1}^{6} Y_t$$

بعد إيجاد قيمة ل F_1 وتحديد قيمة مناسبة أو مبدئية ل α يسير التنبؤ في التمهيد الأسى المفرد كما يلي : في نهاية الفترة $t-1$ يكون التنبؤ بالقيمة في الفترة **t** أي F_t:

$$F_t = \alpha Y_{t-1} + (1-\alpha)F_{t-1}$$

وبمجرد ظهور القيمة الفعلية في الزمن **t** أي Y_t نستفيد منها لتحسين التنبؤ في الزمن $t+1$ والذي يكون :

$$F_{t+1} = \alpha Y_t + (1-\alpha)F_t$$ وهكذا.

مثال (٢ , ٤)

الجدول التالي يوضح المبيعات الشهرية (بوحدات معينة) لنوع من اللمبـات الكهربائية بمتجر كبير في سنة ما ، والتنبؤ بشهر للأمام وأخطـاء التنبـؤ. اسـتخدمت في التنبؤ (اختياراً) $\alpha = 0.02$.

رقم الشهر (t)	المبيعات Y_t	التنبؤ F_t	خطأ التنبؤ e_t
١	١٩	٢٠,٠٠	-1.00
٢	١٩	١٩,٩٨	-0.98
٣	١٨	١٩,٩٦	-1.96
٤	٢٠	١٩,٩٢	$+0.08$
٥	٢٢	١٩,٩٢	$+2.08$
٦	٢٢	١٩,٩٦	$+2.04$
٧	٢٠	٢٠,٠٠	٠
٨	٢٣	٢٠,٠٠	$+3.00$
٩	٢٢	٢٠,١٠	$+1.90$
١٠	٢٥	٢٠,١٠	$+4.90$
١١	٢٤	٢٠,٢	$+3.80$
١٢	٢٤	٢٠,٣	$+3.70$

جدول (٢ , ٤)

العمود الثالث بجدول (٢ , ٤) يعطي التنبؤ لكل شهر من الشهر السـابق لـه. للبـدء في التنبؤ نحتاج لقيمة ل F_1 وقد قدرت هذه القيمة من القيم ال٦ الأولي :

$$F_1 = \frac{19+19+18+20+22+22}{6} = 20$$

للتنبؤ بالقيمة في الشهر ٢ من الشهر ١ نحسب باستخدام (٣ , ٤) وبأخذ $\alpha = 0.02$:

$$F_2 = 0.02 \times 19 + 0.98 \times 20 = 19.98$$

وبظهور القيمة الفعلية في الشهر ٢ أي Y_2 نجد أن الخطأ في هذا التنبؤ كان:

$$e_2 = 19 - 19.98 = -0.98$$

وبالتالي لتحديث هذا التنبؤ ليعطي تنبؤ من الشهر ٢ بالقيمة في الشهر ٣ نحسب:

$$F_3 = 0.02 \times 19 + 0.98 \times 19.98 = 19.96$$

ولأن القيمة الفعلية في ذلك الشهر هي ١٨ فهناك خطأ تنبؤ -1.96

وتستمر عملية التنبؤ على هذا المنوال لنحصل على القيم في العمودين الأخيرين.

إذا أردنا أن نتأكد من أن اختيارنا ل $\alpha = 0.02$ كان سليماً نحاول قيم مختلفة ل α ، نحسب أخطاء التنبؤ في كل حالة وبالتالي متوسط مربعات الخطأ ، وتكون قيمة α التي تعطي أقل متوسط مربعات خطأ هي الأنسب.

فترة ثقة للتنبؤ

يمكن أيضاً إنشاء فترة ثقة تقريبية للتنبؤ بالقيمة في الزمن $t+1$ من الزمن **t** .

فترة الثقة بدرجة ثقة $(1-\beta)\times 100\%$ للتنبؤ F_{t+1} من الزمن **t** نأخذ الشكل:

$$F_{t+1} \pm Z_{\frac{\beta}{2}} 1.25\delta(t+1)$$

حيث :

$$\delta(t+1) = \frac{\sum_{t=1}^{t+1} \left| Y_t - F_t \right|}{t+1}$$

و $Z_{\frac{\beta}{2}}$ القيمة في التوزيع الطبيعي المعياري التي تليها مساحة $\frac{\beta}{2}$.

مثال (٤,٣)

مستخدماً بيانات مثال (٤,٢) أنشئ فترة ثقة بدرجة ثقة ٩٥٪ للتنبؤ في الشهر ٦ . هنا :

$\beta = 0.05$ و $Z_{\frac{\beta}{2}} = 1.96$ من جدول التوزيع الطبيعي .

كذلك من جدول (٢ , ٤) نجد :

$$F_6 + 19.96$$

$$\delta(6) = \frac{1}{6}\left[\left|19-20\right| + \left|19-19.98\right| + \left|18-19.96\right| + \left|20-19.92\right| + \left|22-19.92\right|\right.$$

$$\left.\left|22-19.96\right|\right] = \frac{1}{6}\left[1.0+0.98+1.96+0.08+2.08+2.04\right]$$

$$= \frac{8.14}{6} = 1.36$$

وبالتالي فترة الثقة :

$$19.96 \pm 1.96 \times 1.25 \times 1.36$$

أو 19.96 ± 3.33

أو $(16.63, 23.29)$

وتفسر تلك الفترة بأننا على ثقة قدرها ٩٥٪ أن القيمة الحقيقية المتنبأ بها تراوح بين ١٦,٦٣ و ٢٣,٢٩.

إشارة التتبع Tracking signal

قد يتغير معدل التغير في السلسلة الزمنية مع مرور الوقت ، بحيث يجعلنا نتساءل : هل اختيارنا لقيمة α لا يزال مناسباً ؟ بمعني آخر ، هل لازال التنبؤ باستخدام تلك القيمة ل α يعطي تنبؤات بدرجة معقولة من الصحة ؟.

تساعدنا إشارة التتبع في الإجابة على هذا السؤال.

أفرض أن السلسلة الزمنية عندما استخدمت فيها القيمة α للتنبؤ كانت أخطاء التنبؤ $e_t(\alpha)$ حيث $t = 1,\dots,n$. تعرف إشارة التتبع TS كما يلي :

$$TS = \frac{\sum_{t=1}^{n} e_t(\alpha)}{\sum_{t=1}^{n} \left|e_t(\alpha)\right| \big/ n}$$