المحور السادس: عموميات حول السلاسل الزمنية وتقدير مركباتها

1.1 ماهية السلاسل الزمنية:

في أدبيات الإحصاء يمكن التعرف على ثلاثة أنواع مختلفة من البيانات، هــــــي: البيانات التجريبية وبيانات الحصــر (المســح) والبيانات الزمنية. وتعتمد الفلســفة الخاصــة بالبيانات التجريبية على الأســلوب التجريبي والذي يبدأ بتحديد العوامل الهامة والتي يعتقد الباحث أن لها تأثير معنوي على الظاهرة أو المتغير موضـع الدراسـة، ثم يتم الحصـول على البيانات من خلال تصـميم تجربة تعتمد على مبدأ العشـوائية تسـمح بقياس تأثير أحد أو بعض هذه العوامل على الظاهرة أو المتغير موضع الدراسة فــــــي ظل ثبات العوامل الأخرى. وتعتمد الفلسفة الخاصة ببيانات الحصر أو المسح علـــى مبدأ الحصول على البيانات عن طريق حصــر أو مسـح الوضـع القائم للظواهر موضــع الدراســة كما هو دون محاولة التحكم في العوامل المختلفة التي قد تؤدي إلى الحالة التي توجد عليها هذه الظواهر. أما البيانات الزمنية فيتم الحصول عليها من خلال رصـد البيانات أو القيم التي تعبر عن الظاهرة أو المتغير موضــع الدراســة على فترات زمنية متتالية بهدف تحقيق عدة أهداف أهمها اكتشــــاف نمط التطور التاريخي للظاهرة أو المتغير موضـــــع الدراســــة، وكيفية الاســـتفادة من هذا النمط في التنبؤ بهذه الظاهرة في المســـتقبل. ويطلق على البيانات الزمنية "الســـلاســـل الزمنية" وهي الموضوع الرئيسي لهذا الكتاب.

تعريف: السلسلة الزمنية هي مجموعة من المشاهدات أو القياسات التي تؤخذ على إحــــــــدى الظواهر الاقتصادية – الاجتماعية – الطبية – الطبيعية – ...... على فترات زمنية متتابعة عادة ما تكون متساوية الطول.

ويمكن رصـد السـلاسـل الزمنية في شـتى أنواع المعرفة وميادين التطبيق المختلفة. ففي الاقتصـاد يمكن رصــد بيانات الدخل الســنوي وقيمة التحويلات الخارجية الســنوية والإيداعات ربع الســنوية في أحد البنوك وعدد العاطلين الشـهري وغيرها من البيانات. وفي علم الاجتماع يمكن رصـد عدد الجرائم الأسـبوعي وعدد حالات الطلاق أو الزواج السنوي وغيرها. وفي مجال التعليم يمكن رصد السلاسل الزمنية الخاصة بتطور أعداد الطلبة الســنوي في مراحل التعليم المختلفة وأعداد المدارس والمدرسـين الســنوية في الكليات المختلفة. وفي مجال الطب يمكن رصد السلاسل الزمنية الخاصة بتطور الأمراض المختلفة ومدى التزايد أو التناقص في الإصابة بهذه الأمراض مثل التطور التاريخي لنسبة المصابين بالذبحة الصدرية أو الأورام الخبيثة، كما يمكن رصـد السـلاسـل الزمنية الخاصــة برسـم القلب أو الدماغ. وفي مجال الأرصـاد الجوية يمكن رصـد السلاسل الزمنية الخاصة بكمية الأمطار الشهرية والسلاسل الزمنية الخاصة بسرعة الرياح ونسبة الرطوبة ودرجات الحرارة وغيرها. وفي مجال البيئة يمكن رصد السـلاسـل الزمنية الخاصــة بتطور نسـب التلوث في

الأجواء المحيطة وتطور متوسط الحموضة في مياه الأمطار السنوية ونسب الأكسوجين المذاب فـي الميـاه كمقياس لتلوث المياه. وفي مجال الزراعة يمكن رصـد السـلاسـل الخاصـة بتطور الإنتاج السـنوي من المحاصيل الزراعية والدخل السنوي الناتج من قطاع الزراعة. وفي مجال الكيمياء يمكن رصد درجة الحرارة التي تؤخذ كل دقيقة من عملية كيميائية معينة، وفي مجال الهندسـة يمكن رصـد تطور نسـب الوحدات المعيبة الشهرية وتطور إنتاجية العامل السنوية في أحد المصانع.

وتختلف السلسلة الزمنية عن البيانات التجريبية وبيانات الحصر في ثلاث نقاط أساسية هي:

1. تؤخذ بيانات السـلسـلة الزمنية على فترة زمنية طويلة نسـبيًا يعتقد أنها تؤثر على الظاهرة أو المتغير موضـع الدراسـة، بينما تؤخذ البيانات التجريبية أو بيانات الحصـر (المسـح) عند نقطة زمنية معينة أو على الأكثر في فترة زمنية قصـيرة يعتقد أنها لا تؤثر بشـكل معنوي على الظاهرة أو المتغير موضـع الدراسة. وعادة ما تسم هذه البيانات بالبيانات المتقاطعة cross sectional data
2. يتم دراسـة السـلسـلة الزمنية عادة بمعزل عن العوامل الأخرى – بخلاف الزمن – التي قد تؤثر عليها وعن الظواهر الأخرى التي قد ترتبط معها في علاقة إحصائية.
3. عادة ما تكون بيانات أو مشـاهدات السلسـلة الزمنية مرتبطة ببعضـها البعض، ويأخذ الارتباط بين هذه المشاهدات أشكالاً وأنماطاً عديدة تختلف باختلاف طبيعية الظاهرة، ومن ثم فإن ترتيب المشاهدات في السـلاسـل الزمنية ذو أهمية خاصـة، ولذلك فإن معظم الأسـاليب التي تسـتخدم في تحليل البيانات التجريبية أو بيانات الحصـر لا تكون صـالحة لتحليل السـلاسـل الزمنية، وبالتالي كان لابد من ابتكار وتطوير أدوات وأساليب خاصة لتحليل السلاسل الزمنية.

وقد خصَّص الإحصـاء مجالاً منفردًا لتحليل البيانات الزمنية يعرف بمجال لسلاسل الزمنية والذي تطور تطوراً هائلاً في العقود الثلاثة الأخيرة من القرن العشرين بسبب المنهجية الحديثة التي قدمها العالمان بوكس G.E.Box جينكنز Jenkins ... في سـنة 1976 والتي يمكن اعتبارها بحق البداية الحقيقية لتحليل الحديث للسـلاسـل الزمنية والسـبب الحقيقي وراء القفزات العلمية الهائلة التي حدثت في هذا المجال. وللمزيد من التفاصـيل حول أنواع البيانات الإحصـائية والفروق الأسـاسـية بينهم يمكن للقارئ الرجوع إلى شـعراوي وإسماعيل (2002).

وتعرض السـلسـلة الزمنية عادة في صـورة جدول أو خط أو منحنى بياني يعـرف الخط التاريخي أو المنحنى الزمني time series plot كما في الأمثلة الآتية:

<u>مثال (1):</u> يوضـح الجدول الآتي حجم الإنتاج السـنوي للبترول بالمليون متر مكعب في إحدى الدول من سنة 1990 إلى سنة 1997:

السنة	1990	1991	1992	1993	1994	1995	1996	1997
حجم الإنتاج	5210	5820	5655	7800	8100	8010	9200	9335

ويمكن عرض هذه البيانات في شكل خط بياني (تاريخي) في شكل (1):

شكل (1): الخط البياني (التاريخي) لبيانات المثال (1)

مثال (2): يوضـح الجدول الآتي تطور قيمة الإيداعات ربع السـنوية بالمليون دولار في أحد البنوك في سنتي 1998 و 1999:

السنة	الموسم (الفصل)	الإيداعات
1998	1	42
	2	46
	3	50
	4	56
1999	1	45
	2	49
	3	54
	4	60

ويمكن عرض هذه البيانات في شكل (2)

شكل (2): العرض البياني لبيانات المثال(2)

1.2 أنواع السلاسل الزمنية:

عند دراسة السلاسل الزمنية لبعض الظواهر قد يكون من الممكن أخذ قياسات أو قراءات عند كل لحظة زمنية ويقال لهذه السلاسل بأنها سلاسل متصلة continuous، ومن أمثلة هذه السلاسل درجات الحرارة ورسم القلب ورسم الدماغ. أما معظم السلاسل الزمنية التي تنشأ في الواقع فتتكون من قراءات أو مشاهدات مأخوذة عند فترات زمنية محددة مسبقا. وقد تكون هذه الفترات دقائق أو ساعات أو أيام أو أسابيع أو شهور أو سنوات. وتعرف هذه السلاسل بالسلاسل المتقطعة discrete time series بغض النظر عن طبيعة الظاهرة أو المتغير موضع الدراسة، أمثلة هذه السلاسل الدخل القومي السنوي وسعر الإقفال اليومي لأحد الأسهم في بورصة الأوراق المالية وعدد الحوادث الأسبوعية التي تحدث على أحد الطرق وعدد خريجي إحدى الكليات السنوي وكمية الأمطار الشهرية والسلاسل الزمنية المتقطعة السلاسل التي سنتعامل معها فقط في هذا الكتاب، أي أننا سنفترض دائمًا أن السلسلة متاحة فقط عند نقاط زمنية متقطعة تبعد عن بعضها فجوات زمنية متساوية الطول.

وفي الواقع يمكن الحصول على السلسلة الزمنية المتقطعة بمعاينة سلسلة زمنية متصلة وذلك بأن يتم رصد أو تسجيل القراءات فقط عند نقاط زمنية محدَّدة متساوية الأبعاد، كما يمكن الحصول على السلسلة

الزمنية المتقطعة بتراكم متغير معين خلال فترة زمنية مثل كمية الأمطار التي عادة ما تتراكم خلال يوم أو شهر مثلاً أو الناتج السنوي من أحد المحاصيل الزراعية والاهتمام الأساسي لهذا الكتاب هو كيفية بنـــــاء النماذج للسلاسل الزمنية المتقطعة واستخدام هذه النماذج في التنبؤ بالمشاهدات المستقبلية.

1.3 أهداف دراسة السلاسل الزمنية:

تدرس السلاسل الزمنية عادة لتحقيق عدد من الأهداف. وقد يكون أول أهداف هذه الدراسة هو استخدام السلسلة الزمنية لوصف وتصوير المعلومات المتاحة عن فترة زمنية توضح تطور الظاهرة المدروسة أي وصف الملامح والسمات الرئيسية للسلسلة. ويساعد وصف السلسلة إلى حد كبير في تحديد النموذج الذي يمكن أن يكون مناسبًا لتحقيق الأهداف الأخرى والتعرف على حركات الصعود والهبوط في السلسلة الزمنية والتعرف على المكونات الرئيسية مثل الاتجاه العام والتغيرات الموسمية كمـا سنرى في نهاية هذا الباب. أما الهدف الثاني من دراسة السلاسل الزمنية فهو التفسير، ويقصد به توضيح وشرح التغيرات التي تحدث في الظاهرة باستخدام السـلاسـل الزمنية الأخرى التي ترتبط بها أو باستخدام عوامل البيئة المحيطة بالظاهرة ومثال ذلك تفسير التغيرات التي تحدث في سلسلة المبيعات الخاصة بإحدى السلع باستخدام السلسلة الخاصة بتغيرات أسعار هذه السلعة أو بمعرفة القرارات الاقتصادية التـــي اتخذت وكانت لها علاقة مباشرة على التطور التاريخي للظاهرة أو تفسير التغيرات التي تحدث في سلسلة عدد الحوادث التي تحدث على طريق جدة/ مكة المكرمة بالسلسلة الزمنية لعدد الحجاج السنوي أو السلسلة الزمنية لعدد المعتمرين الشهري أو الإجراءات الأمنية التي اتخذت للحد من هذه الحوادث وربط التغيرات التي تحدث في السلسلة موضع الدراسة بالتغيرات التي تحدث في السلاسل والعوامل المحيطة من شأنه فهم آلية عمل السلسلة وتفسير الأنماط والتغيرات المنتظمة وغير المنتظمة التي تتعرض لها الظاهرة موضع الدراسة ومدى تأثير كل منها عليها. والهدف الثالث من دراسة السـلاسـل الزمنية هو الرقابة والتحكم، فقد تستخدم الخرائط الزمنية في مراقبة جودة الإنتاج وذلك من أجل التحكم في مستوى كفاءة العملية الإنتاجية وذلك باتخاذ القرارات المناسبة من وقف العملية الإنتاجية وتعديل مسارها أو استمرارها.

أما أهم أهداف دراسة السلاسل الزمنية على الإطلاق فهو التنبؤ بالمشاهدات المستقبلية والذي عادة ما يمثل الهدف النهائي من تحليل السلاسل الزمنية. وهذا الهدف هو أوضح الأهداف وأكثرها شعبية بالنسبة لدارس الإحصـــاء أو مســـتخدمه والذي من أجله كتبت عشـــرات الكتب العالمية والآلاف من الأبحاث المتخصِّصة. فتحليل السلاسل الزمنية يبدأ عادة بالتعرف على النمط المناسب لشرح آلية تطور هذه السلسلة واستكمال هذا النمط مستقبلاً. والفرض الأساسي في أساليب التنبؤ المستخدمة هو أن هذا النمط الذي تم التعرف عليه سيستمر في المستقبل القريب، وتجدر الإشارة إلى أنه لا يمكن لأي أسلوب تنبؤ أن يعطي

نتائج جيدة إذا لم يستمر هذا النمط، ولذلك فإنه ينصح دائمًا بالتنبؤ بالقيم المستقبلية القريبة وتحديثها بمجرد الحصول على أي مشاهدة جديدة.

1.4 قياس أخطاء التنبؤ:

عادة ما تدرس السلسلة الزمنية بغرض اكتشاف نمط التطور التاريخي للظاهرة واستغلال هذا النمط في التنبؤ بالقيم المســـتقبلية. وأي تنبؤ مســـتقبلي لأي ظاهرة لابد أن يحتوي على قدر معين من عدم التأكد، ويمكن ترجمة هذه الحقيقة بإدراج مركبة خطأ هي error component في نموذج التنبؤ. ومركبة الخطأ هي المركبة غير النمطية التي تعبر عن العوامل التي لا يمكن شرحها باستخدام التغيرات النمطية أو المنتظمة في السلسلة. وكلما كانت هذه المركبة صغيرة زادت قدرتنا على التنبؤ والعكس صحيح. إذا افترضنا أن قيمة الظاهرة موضـــع الدراســـة عند الزمن t هي y_1 وأن التنبؤ بالظاهرة عند الزمن t هو $\hat{y}_1$، فإن الخطأ في التنبؤ عند الزمن t يعرف كالآتي: $e_1 = y_1 - \hat{y}_1 \quad ; \quad t = 1,\ 2, \ldots \ldots, n$

حيث يرمز n إلى طول السلسلة أي عدد المشاهدات.

وفحص أخطاء التنبؤ المتتالية، يوضـــح مدى ملاءمة أسـلوب التنبؤ المســـتخدم، فكما هو معروف من دراسة الانحدار أن أسلوب التنبؤ الملائم لابد أن ينتج أخطاءً تتصف بطابع العشوائية، أي أخطاء خالية من أي تغيرات منتظمة – كما في شكل (3) – بالإضافة إلى بعض الشروط الأخرى. وإذا كانت هذه الأخطاء محتملة بحيث يمكن اعتبار أسلوب التنبؤ ملائم فإنه يجب قياس حجم هذه الأخطاء لتقدير دقة التنبؤ.

<u>شكل (3):</u> أخطاء عشوائية

ولقد عرف الفكر الإحصائي طرقا عديدة لقياس حجم الأخطاء أهمها ما يلي:

1- مجموع الأخطاء sum of errors ويرمز له عادة بالرمزSE ويعرف على الصورة الآتية:

$$SE = \sum_{t=1}^{n} e_1 = \sum_{t=1}^{n} (y_1 - \hat{y}_1)$$

وهذا المقياس لا يفيد كثيرًا، حيث أنه من المعروف أنه إذا كانت الأخطاء عشـــوائية فإن هذا المجموع عادة ما يكون قريبًا جدًا من الصفر بغض النظر عن حجم هذه الأخطاء.

2- متوسـط الانحرافات المطلقة mean absolute deviation والذي يرمز له عادة بالرمز MAD ويعرف على الصورة الآتية: $MAD = \frac{1}{n}\sum_{t=1}^{n}|e_1 = \sum_{t=1}^{n}\ |y_1 - \hat{y}_1|$

وبالرغم من معقولية هذا المقياس إلا أنه لا يستخدم كثيرًا في مجالات السلاسل الزمنية نظرا لصعوبة خصائصه الإحصائية.

3− متوسـط مربعات الأخطاء mean squared error ويرمز له عادة بالرمزMSE ويعرف على الصـورة الآتية:

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 = \frac{1}{n}\sum_{t=1}^{n}(y_1 - \hat{y}_1)^2$$

ويلاحظ أن طريقة المربعات الصغرى least squares method المعروفة في مجالات الانحدار والسلاسل الزمنية تعتمد على تصغير مجموع مربعات الأخطاء SSE أو تصغير متوسط مربعات الأخطاء MSE وذلك لأن المقام n والذي يمثل عدد الوحدات الزمنية المتاحة (عدد المشـاهدات) هو مقدار ثابت وبصـفة عامة يمكن القول بأن خصائص هذا المقياس الإحصائية أسهل كثيراً من خصائص متوسط الأخطـاء المطلقة MAD

4− متوسـط الأخطاء النسـبية المطلقة mean absolute percentage error والذي يرمز له عادة بالرمز MAPE ويعرف على الصورة التالية:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_1 - \hat{y}_1}{y_1}\right| = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{e_1}{y_1}\right|$$

ويتميز هذا المقياس عن كل المقاييس بأنه مقياس نسـبي، أي لا يعتمد على الوحدات المسـتخدمة ولكن خصائصه الإحصائية أصعب من خصائص متوسط مربعات الأخطاء *MSE* ولذلك فإن هذا المقياس عادة ما يستخدم في الدراسات الوصفية التي لا تستدعي القيام باستدلالات إحصائية.

1.5 اختيار أسلوب التنبؤ المناسب:

من أهم عناصر تحليل السلاسل الزمنية اختيار أسلوب التنبؤ المناسب. واختيار أسلوب التنبؤ المناسب ليس بالعمل الهين وإنما هو عمل صـعب وشـاق ويحتاج من الإحصـائي ومتخذ القرارات التحلي بالصـبر وعدم اليأس بالإضـافة إلى مقومات العمل الأسـاسـية من علم وخبرة ومهارة. ويعتمد الإحصـائي أو متخذ القرارات بصـفة عامة في اختياره لأسـلوب التنبؤ المناسـب على بعض المعايير أو العوامل العامة أهمها ما يلي:

1. تصـغير حجم أخطاء التنبؤ أول هذه المعايير التي يجب أن يضـعها الإحصـائي أو متخذ القرارات نصـب عينيه عند اختياره أسـلوب التنبؤ، ومن ثم نموذج التنبؤ المناسـب، وعادة ما يقاس حجم هذه الأخطاء بأحد المقاييس الثلاثة التي سبق ذكرها (MAD- MSE- MAPE).
2. نوعية التنبؤ المطلوب، فإذا كان تنبؤ النقطة point forecast هو المطلوب من الدراسـة، فإن استخدام أحد الأسـاليب أو النماذج التقليدية البسـيطة قد يكون كافيا لتحقيق هذا الهدف. وفي الكثير من

الدراســـــات قد يكون تنبؤ الفترة forecast interval هام وكذلك اختبارات الفروض، وفي مثل هذه الحالات لابد من استخدام أسلوب تنبؤ حديث أكثر دقة وتنظيمًا مثل أسلوب بوكس وجينكنز.

3. عدد المشاهدات المتاحة، فإذا كان عدد المشاهدات صغيرًا فإن استخدام أحـد الأساليب الحديثة ليس له ما يبرره ويفضل استخدام أحد الأساليب التقليدية.
4. تكاليف أسلوب التنبؤ ومدى توافر البرامج الإحصائية ذات الصلة.
5. سهولة العمليات الإحصائية والحسابية الضرورية وفهم أسلوب التنبؤ المستخدم.
6. مدى تحقق الفروض النظرية التي يعتمد عليها أسـلوب أو نموذج التنبؤ المناسـب وهو أهم المعايير التي يجب أن تؤخذ في الاعتبار عند اختيار مثل هذا الأسلوب.

مما سبق يتضح للقارئ بأن أفضل أسلوب للتنبؤ ليس بالضرورة هو الأسلوب الذي يحقق أعلى دقة أو أصغر حجم أخطاء ممكن، فقد يستخدم أحد الأساليب بسبب نوعية التنبؤ المطلوب، وقد يستخدم أسلوب آخر بسبب صغر عدد المشاهدات المتاحة، وقد يستخدم أسلوبًا ثالثًا بسبب انخفاض تكاليفه، وقد يستخدم أسلوب رابع بسبب سهولة عملياته الإحصائية والحسابية، وقد يستخدم أسلوبًا خامسًا لأن الفروض النظرية التي يعتمد عليها تتوافق مع بيانات السلسـلة المتاحة. وعادة ما يعتمد أسـلوب التنبؤ المستخدم على قدرة الإحصـــائي أو متخذ القرارات في تحقيق التوازن لكل هذه المعايير. وبصـــفة عامة يمكن القول بأن طريقة التنبؤ التي يجب استخدامها هي أسهل وأبسط طريقة يمكن تنفيذها في الزمن المتاح والتي تفي باحتياجات وظروف التنبؤ بأقل تكاليف ممكنة، وللمزيد من التفاصيل حول قياس الأخطاء ومعايير التنبؤ يمكن للقارئ الرجوع إلى Bowerman and O'Connell (1987) أو إلى (1994) Gaynor Kirkpatrick and

1.6 طرق التنبؤ:

يمكن تجميع طرق التنبؤ الكمية المعروفة في أدبيات السلاسل الزمنية في أسلوبين أساسيين هما.

أ– أسلوب الانحدار regression approach, ويعتمد هذا الأسلوب على تحديد المتغيرات الأخرى التي قد ترتبط بعلاقة سببية بالظاهرة أو المتغير موضـع الدراسـة الذي يراد التنبؤ به والذي عادة ما يعرف بالمتغير التابع dependent variable – ثم تحديد النموذج الإحصائي أو العلاقة الدالية الملائمة التي توضح الكيفية التي يرتبط بها هذا المتغير بالمتغيرات الأخرى والتي تأخذ في العرف الإحصـــائي أســـماء عديدة مثل المتغيرات المســـتقلة independent variables أو المتغيرات المفســـرة regressors أو المتغيرات المنبئـة predictors. وباسـتخدام هذا النموذج يمكن التنبؤ بالمتغير التابع موضـع الدراسـة في المسـتقبل إذا أمكن تحديد أو معرفة القيم المستقبلية للمتغيرات المفسـرة. وتعرف النماذج التي تندرج تحت مظلة هذا الأسـلوب أحيانا بالنماذج السـببية causal models ويستخدم هذا الأسـلوب في كافة أنواع المعرفة ومجالات التطبيق

الخاصة الاقتصادية والاجتماعية والبيئية منها حيث يسمح هذا الأسلوب بتقييم أثر المتغيرات المتضمنة والتي عادة ما تعكس أثر الأنظمة والسياسات والقرارات المختلفة، فعلى سبيل المثال قد نستطيع تفسير السلسلة الخاصة بقيمة المبيعات اليومية لإحدى السلع جزئيًا بواسطة سلسلة أسعار هذه السلعة والسلاسل الخاصة بالدخل الفردي وأسعار السلع البديلة. وهذا الأسلوب بالرغم من شعبيته يعاني من بعض العيوب أهما ما يلي:

1- صعوبة تحديد المتغيرات المفسرة التي ترتبط بالمتغير التابع أو الظاهرة موضع الدراسة.

2- تطبيق هذا الأسلوب يتطلب توافر بيانات تاريخية تفصيلية عن جميع المتغيرات المفسرة والقدرة على معرفة قيم هذه المتغيرات – أو على الأقل التنبؤ بها- عند الأزمنة التي يراد التنبؤ بالظاهرة عندها.

3- يفترض عدم الارتباط بين مشاهدات المتغير أو الظاهرة موضع التنبؤ، وهو فرض غير واقعي ولا يتفق مع مفهوم السلسلة الزمنية باعتبارها مجموعة من المشاهدات المرتبطة وعادة ما يؤدي هذا الفرض غير الواقعي إلى تنبؤات غير موثوق بها.

وعادة ما يتوافر لدى الباحث مشاهدات تاريخية عن المتغير موضع الدراسة فقط ويريد التنبؤ بالمشاهدات المستقبلية لهذا المتغير بالاعتماد فقط على هذه المشاهدات. في مثل هذه الحالات يستخدم الأسلوب الثاني للتنبؤ التالي.

ب – تحليل السلاسل الزمنية time series analysis والذي يضم تحت مظلته ما يعرف بنماذج السلاسل الزمنية، ويعتمد هذا الأسلوب على تحليل البيانات التاريخية التي أخذت عن الظاهرة أو المتغير موضع الدراسة وذلك بغرض تحديد نمط البيانات. بعد ذلك – وبافتراض أن هذا النمط سيستمر في المستقبل – يستكمل هذا النمط لإعطاء التنبؤات المطلوبة. فعلى سبيل المثال إذا كان الهدف من الدراسة التنبؤ بعدد الحجاج السنوي فقد يستطيع الباحث تفسير سلوك هذا المتغير جزئيا بواسطة عدد السكان في الدول الإسلامية ومتوسط دخل الفرد في هذه البلاد ومتوسط التكاليف المطلوبة لأداء هذه الفريضة، ولكن جزء كبير من تطور عدد الحجاج قد يعود إلى بعض العوامل الأخرى التي لا يمكن أخذها في الاعتبار بسهولة مثل الوازع الديني وحالة الطقس وغيرها من العوامل التي يكون من الصعب أو المستحيل إدراجها في النماذج السببية بسهولة. وفي هذه الحالة قد يفضل دراسة التطور التاريخي لعدد الحجاج السنوي بمعزل عن جميع العوامل المفسرة الأخرى واكتشاف الكيفية التي يتطور بها عدد الحجاج واستخدام أحد نماذج السلاسل الزمنية لاستكمال هذه السلسلة في المستقبل.

وأسلوب السلاسل الزمنية والذي يضم ما يعرف بنماذج أو طرق السلاسل الزمنية هو محور الاهتمام الرئيسي لهذا الكتاب، ولن نتعرض للنماذج السببية بأي حال من الأحوال. والسؤال الآن هو كيف يمكن

التنبؤ بالظاهرة أو المتغير موضع الدراسة باستخدام أسلوب السلاسل الزمنية دون اللجوء إلى أي متغيرات أخرى مفسرة؟ للإجابة عن هذا السؤال يمكن القول بأن أدبيات السلاسل الزمنية قد عرفت العديد من الطرق ونماذج السلاسل الزمنية والتي يمكن تقسيمها إلى ثلاثة أنواع رئيسية هي:

- النماذج المحددة (غير العشوائية) deterministic models
- والطرق الحسية ad hoc methods
- ونماذج السلاسل الزمنية العشوائية stochastic time series models

ونقدم فيما يلي عرضًا سريعًا لهذه الطرق والنماذج.

1.6.1 النماذج المحددة Deterministic Models

نعرف من دراستنا في علم الإحصاء أن نموذج المتوسط mean model يمكن التعبير عنه في الصورة العامة

$$y_1 = E(y_1) + \varepsilon_1$$

حيث ε_1 متغيرات عشوائية غير مرتبطة توقعها الصفر وتباينها ثابت ويقال أن هذا النموذج محدد deterministic أو غير عشوائي nonstochastic إذا أمكن التعبير عن $E(y_1)$ كدالة رياضية مباشرة في الزمن t ولتكن $f(t,\beta)$ حيث يرمز المتجه β إلى معالم هذه الدالة الرياضية. وفي هذه الحالة يمكن التعبير عن مشاهدات السلسلة الزمنية y_1 على الصورة:

$$y_1 = f(t,\beta) + \varepsilon_1 \quad ; \quad t = 1,2,\dots,n \qquad (1.6.2)$$

ويعني هذا أن المشاهدات المستقبلية للسلسلة يمكن التعبير عنها على الصورة:

$$y_h = f(h,\beta) \quad ; \quad h = t+1\,, \qquad t+2, \dots\dots$$

أي أن هذه النماذج تفترض أن المشاهدات المستقبلية للسلسلة تأخذ شكلاً رياضيًا محددًا، أي غير عشوائي

وتعتمد النماذج المحددة (1.6.2) على فرضيين أساسيين. الفرض الأول أن الدالة $f(t,\beta)$ دالة رياضية محددة ليس لها طابع العشوائية، والفرض الثاني أن ε_1 متغيرات عشوائية غير مرتبطة توقعها الصفر وتباينها ثابت وتؤدي هذه الفروض إلى أن المتغيرات $y_1, y_2, \dots, y_n$ تكون متغيرات عشوائية غير مرتبطة، ومن أمثلة الدوال الرياضية التي تستخدم في هذه النماذج كثيرات الحدود والدوال الأسية والدوال المثلثية، وفيما يلي عرضا مبسطا لكثيرات الحدود والدوال الأسية.

كثيرات الحدود Polynomials:

ويفترض هنا أن الدالة f(t) أي متوسـط الظاهرة تأخذ إحدى صـور كثيرات الحدود في الزمن t وتعتبر الصورة الخطية أهم هذه الصور وتعرف على الشكل الآتي:

$$E(y)_1 = f(t) = \beta_0 + \beta_1 t$$

وتكون هذه الدالة ملائمة إذا أمكن تمثيل متوسط الظاهرة بواسطة خط مستقيم وذلك بعد توقيع مشاهدات السلسلة على ورقة الرسم البياني كما في شكل (4) .

وتفترض كثيرة الحدود الخطية أن متوسـط الظاهرة يتزايد بمعدل ثابت β_1 كما في الشـكل (4.a) أو يتناقص بمعدل ثابت β_1 كما في الشكل (4.b)

ويمكن إيجاد تقديري المربعات الصغرى $\hat{\beta}_1, \hat{\beta}_2$ لمعلمتـي النموذج β_0, β_1 بإجراء انحدار السلسلة y_1 على الزمن t، ومن ثم يمكن التنبؤ بـالمشـاهـدات المسـتقبلية بـاسـتخـدام النموذج المقـدر الآتي: $\hat{y}_1 = \hat{\beta}_1 + \hat{\beta}_2$

وفي بعض الأحيان قد يكون الشـكل الخطي غير ملائم لتمثيل متوسـط الظاهرة أو الدالة $f(t)$ ومن الأفضل تمثيل هذا المتوسط بكثيرة حدود من الدرجة الثانية كما فـي الشكل (5)− والتي تأخذ الصورة الآتية:

$$E(y_1) = f(t) = \beta_0 + \beta_1 t + \beta_2 t^2$$

وتستخدم طريقة المربعات الصغرى العادية لإيجاد التقديرات $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ وذلك بإجراء الانحدار الخطي المتعدد لقيم السلسلة y_1 على المتغيرين t, t^2. وسنوضح هذا بالتفصيل عند قياس الاتجاه العام باستخدام تحليل الانحدار.

<u>شكل (5)</u>: كثيرة حدود من الدرجة الثانية

النمو الأسي Exponential Growth

في بعض الأحيان قد يفضل تمثيل متوسط الظاهرة موضع الدراسة في شكل دالة أسية – كما في الشكل (6)–على الصورة:

$$E(y_1) = f(t) = ce^{rt} \qquad (1.6.3)$$

حيث c و r مقداران ثابتان يمثلان معلمتي النموذج.

شكل (6): النمو الأسي

ويفترض النموذج الأسى أن متوسط الظاهرة ينمو بنسبة ثابتة وذلك لأن من (1.6.3):

$$\frac{E(y_1)}{E(y_{t-1})} = \frac{ce^{rt}}{ce^{r(t-1)}} = e^r$$

حيث e^r ثابت لا يعتمد على الزمن، ويعني هذا أن: $E(y_1) = e^r E(y_{t-1})$

ويمكن تحويل الصورة (1.6.3) إلى الصورة الخطية بأخذ اللوغاريتمات الطبيعية للطرفين كما يلي:

$\ln f(t) = \ln c + rt$ أي أن $f^*(t) = c^* + rt$ حيث: $f^*(t) = \ln f(t)\ ;\ c^* = \ln c$

وبالتالي يمكن إجراء انحدار لوغاريتم البيانات الأصـلية على الزمن t وإيجاد تقديري المربعات الصـغرى للمعلمتين c^*, r ومن ثم يمكن إيجاد تقدير الثابتc من العلاقة: $c = e^{c^*}$

ومن ثم يمكن التنبؤ بالمشاهدات المستقبلية باستخدام النموذج المقدر الآتي: $\hat{y}_1 = \hat{c}\ e^{\hat{r}t}$

وتعاني الطرق أو النماذج المحددة في تحليل السلاسل الزمنية من العديد من العيوب أهمها:

1– تركز هذه الطرق على المنطق الرياضي في محاولة لإيجاد دالة رياضية جيدة ممكن أن تستخدم في توفيق البيانات أكثر من اهتمامها بمحاولة استكشاف الخصائص الإحصائية الهامة للسلسلة وأهمها نمط الارتباط الموجود بين مشاهدات السلسلة، فهذه النماذج لا تصف خصائص السلسلة الإحصائية ولكنها مجرد نماذج تنتج المشاهدات $y_1, y_2, \ldots . y_n$

2– تفترض هذه الطرق أن التطور طويل الأجل في السلسلة يكون نمطي أو منتظم الشكل وممكن التنبؤ به بشكل كبير.

3– وتفترض أيضًـا هذه الطرق عدم وجود ارتباط بين مشـاهدات السـلسـلة، وهذا الفرض من النادر أن يكون متحققا في مجالات التطبيق المختلفة. بسبب كــل هذه العيوب فإن هذه الطرق عادة ما تؤدي إلى تنبؤات غير دقيقة من الناحية الإحصائية.

وبالرغم من الانتقادات التي توجه إلى هذه الطرق فإن لها بـاع طويل في موضـــــوعات التنبؤ ومازالت تستخدم بكثرة في كافة مجالات التطبيق في البلاد النامية خاصـة في الاقتصـاد والإدارة والبيئة لأنها وسـائل

بسـيطة وغير مكلفة ولا تحتاج إلى خبرات أو مهارات خاصـة من قبل الباحثين أو الدارسـين أو متخذي القرارات.

1.6.2 الطرق الحسية Ad Hoc Methods:

ذكرنا أن طرق التنبؤ المحددة تعتمد على التعبير عن قيمة السلسلة عند الزمن t أي y_1 كدالة رياضية مباشـرة في الزمن. وهذا الاتجاه له عيوبه كما ذكرنا وأهمها أنه يفترض عدم وجود علاقة بين مشـاهدات السلسلة $y_1, y_2, \ldots . y_n$, الاتجاه الثاني في التنبؤ يعتمد على التعبير عن تنبؤ السلسلة عند الزمن t بدلالة حاضر السلسلة y_1 وماضـي السلسلة $y_1, y_2, \ldots . y_{t-1}$,. فإذا افترضنا أن t تمثل نقطة أصل معينة وأننا نريد التنبؤ بقيمة السلسلة بعد k من الفترات الزمنية فإن الاتجاه الثاني يفترض العلاقة الدالية الآتية:

$$\hat{y}_{l+k} = f(y_1, y_2, \ldots, y_{t-1}, y_t) \qquad (1.6.4)$$

ويوجد العديد من الطرق والنماذج التي تنتمي بشـكل أو بآخر للصـورة (1.6.4) وتعتمد على الحس الإنسـاني أكثر من اعتمادها على أسـلوب إحصـائي منظم. ومن أمثلة ذلك طريقة التنبؤ السـطحي وتنبؤ التغير الثابت وطريقة المتوسطات المتحركة البسيطة وطريقة التمهيد الأسي، وفيما يلي نقدم عرضًا مبسطا لهذه الطرق.

التنبؤ السطحي Naive Forecasting:

تستخدم طريقة التنبؤ السطحي قيمة المشاهدة الحالية كتنبؤ مباشر للمشاهدة التالية، أي أن:

$$\hat{y}_{t+1} = y_1 \qquad (1.6.5)$$

والنموذج (1.6.5) يكون ملائمًا عندما تكون قيم السلسلة ثابتة بشكل تقريبي على الفترة الزمنية موضـع الدراسـة، ويحدث هذا عندما يغلب على السلسلة الزمنية محل الدراسـة الطابع غير النمطي (غير المنتظم) أي عندما تتغير السلسلة بشكل عشوائي كبيـر لا يتبع نمطا أو نظامًا أو اتجاها معينًا يمكن معه التنبؤ بقيمة السلسلة في الفترة الزمنية التالية، وتعتبر أسعار الأوراق المالية في البورصـة أشـهر مثال على هذا النـوع مـن السلاسل الزمنية.

والجدير بالذكر أن النموذج السطحي يعطي تنبؤات متحيزة إلى أسفل إذا كانت السلسلة تتزايد باستمرار وذلك لأن التنبؤ $\hat{y}_1$ يكون دائما أقل من القيمة، y_1 والعكس صـحيح، أي أن هذا النموذج يعطي تنبؤات متحيزة إلى أعلى إذا كانت السلسلة تتناقص باستمرار لأن التنبؤ $\hat{y}_1$ يكون أكبر من القيمة y_1 ، ولذلك لا ينصح باستخدام طريقة التنبؤ السطحي في مثل هذه الحالات.

تنبؤ التغير الثابت Constant Change Forecasting:

في الكثير من التطبيقات خاصـة الاقتصـادية منها تتميز بعض السـلاسـل بثبات في التغيرات المتتالية، فإذا افترضنا أن t تمثل نقطة أصل معينة فإن التغير السابق في السلسلة يكون: $\Delta y_t = y_t - y_{t-1}$

وإذا كانت $\hat{y}_{t+l}$ تمثل التنبؤ بقيمة السلسلة الزمنية، فإن التغير القادم في السلسلة يقدر كالتالي:

$$\hat{\Delta} y_{1+1} = \hat{y}_{t+1} - y_1$$

وبمساواة التغير السابق Δy_1، بالتغير اللاحق $\hat{\Delta} y_{1+1}$ نصل إلى: $\hat{y}_{t+1} - y_t = y_t - y_{t-1}$

ومن ثم نصل إلى النموذج الآتي:

$$\hat{y}_{t+1} = y_t + (y_t - y_{t-1}) \qquad (1.6.6)$$

أي أن التنبؤ في الفترة الزمنية القادمة يسـاوي القيمة الحاضـرة y_1التغير الذي مضـافا إليه قيمة التغير الذي حدث في الفترة السابقة Δy_t.

المتوسطات المتحركة البسيطة Simple Moving Averages:

يعتمد النموذج السـطحي على القيمة الحالية y_t فقط للتنبؤ بالقيمة التالية y_{t+1} ، بينما يعتمد تنبؤ التغير الثابت على أحدث قيمتين y_{t-1}, y_t للتنبؤ بالقيمة التالية y_{t-1}... أما طريقة المتوسطات المتحركة البسـيطـة فتسـتخـدم أحـدث k قيمـة للسـلسـلـة للتنبؤ بـالقيمـة التـاليـة أي تسـتخـدم القيم $y_t, y_{t-1}, y_{t-2}, \dots, y_{t-(k-2)}$, وذلك بأخذ متوسط هذه القيم كما يلي:

$$\hat{y}_{t+1} = \frac{1}{k}\left[y_t + y_{t-1} + \cdots + y_{t-(k-2)} + y_{t-(h-1)}\right] \; ; \; t = k, k+1, \dots . n \quad (1.6.7)$$

وهذا يعني أن: $\hat{y}_{t+2} = \frac{1}{k}\left[y_{t+1} + y_1 + y_{t-1} + \cdots + y_{t-(k-2)}\right]$

أي أنه لإيجاد المتوسط المتحرك البسيط $\hat{y}_{t+2}$ نستخدم نفس القيم التي استخدمت في حساب المتوسط السـابق لـه مباشـرة $\hat{y}_{t+1}$ بعد إحلال القيمة الأحدث y_{t+1} مكان القيمة الأقدم $y_{t-(h-1)}$ وهذا معنى التحرك أي أن المتوسـط يتم تحديثه دائمًا بحذف المشـاهدة الأقدم ووضـع بدلاً منها المشـاهدة التالية. فعلى سـبيل المثال إذا كانت $k = 3$ فإنه يمكن تكوين $(n-3)$ متوسـط متحرك بسـيط مناظر لقيم السـلسـلة المتاحة كما يلى:

$$\hat{y}_4 = \frac{1}{3}[y_3 + y_2 + y_1]$$

$$\hat{y}_5 = \frac{1}{3}[y_4 + y_3 + y_2]$$

$$\hat{y}_6 = \frac{1}{3}[y_5 + y_4 + y_3]$$

$$\vdots \qquad\qquad \vdots$$

$$\hat{y}_n = \frac{1}{3}[y_{n-1} + y_{n-2} + y_{n-3}]$$

واختيار العدد الصحيح k يعتمد على رأي الباحث وخبرته العملية وهو أحد المشاكل التي تواجه مستخدم طريقة المتوسطات المتحركة البسيطة. ودقة التنبؤ تعتمد على اختيار العدد الملائم ولذلك يمكن اختيار هذا العدد بطريقة التجربة والخطأ، حيث تحسب جميع التنبؤات التي تناظر كل قيمة من قيم k الممكنة $(k = 2, 3, \dots, n-1)$ وحساب الأخطاء ومن ثم حساب أحد المعايير الهامة لقياس حجم الأخطاء – وليكن متوسط مربعات –الأخطاء المناظر لكل قيمة من قيم k واختيار قيمة k التي تناظر أصغر قيمة لهذا المعيار. وبالرغم من المشاكل التي قد تواجه الباحث عند اختيار قيمة k الملائمة إلا أن العيب الرئيسي لطريقة المتوسطات المتحركة البسيطة هو إعطاء أوزان متساوية لكل المشاهدات المستخدمة في حساب المتوسط فإذا كانت $k = 8$ مثلاً فإن الوزن الذي يعطى للقيمة الحديثة y_t يساوي الوزن الذي يعطى للقيمة الأقدم y_{t-7}. وهذا عادة ما يتعارض مع خصائص السلاسل الزمنية حيث نميل إلى إعطاء المشاهدات الأحدث أوزانا أكبر . ولذلك يفضل استخدام طريقة المتوسطات المتحركة البسيطة عندما يغلب الطابع العشوائي على بيانات السلسلة.

<u>مثال (3):</u> الجدول الآتي يوضح قيمة المبيعات السنوية من إحدى السلع بملايين الدولارات في الفترة من السنة 1990 إلى سنة 1998:

السنة	1990	1991	1992	1993	1994	1995	1996	1997	1998
قيمة المبيعات	9	11	10	12	11	9	13	11	9

a. استخدام طريقة المتوسطات المتحركة البسيطة في إيجاد جميع التنبؤات الممكنة مرة باستخدام $k = 2$ ومرة باستخدام $k = 3$ وأوجد متوسط مربعات الأخطاء المناظرة في كل حالة.

b. تنبأ بقيمة المبيعات في سنة 1999.

<u>الحل:</u>

		K			
		2		3	
السنة	y_t	$\hat{y}_t$	e_1^2	$\hat{y}_t$	e_1^2
1990	9	−	−	−	−
1991	11	−	−	−	−
1992	10	$\frac{1}{2}[9 + 11] = 10$	0	−	−
1993	12	$\frac{1}{2}[10 + 11] = 10.5$	252.	$\frac{1}{3}[10 + 1 + 9] = 10$	4

1994	11	$\frac{1}{2}[12+10]=11$	0	$\frac{1}{3}[12+10+11]=11$	0
1995	9	$\frac{1}{2}[11+12]=11.5$	6.25	$\frac{1}{3}[11+12+10]=11$	4
1996	13	$\frac{1}{2}[9+11]=10$	9	$\frac{1}{3}[9+11+12]=10.67$	5.429
1997	11	$\frac{1}{2}[13+9]=11$	0	$\frac{1}{3}[13+9+11]=11$	0
1998	9	$\frac{1}{2}[11+13]=12$	9	$\frac{1}{3}[11+13+9]=11$	4
1199	؟	$0\frac{1}{2}[9+11]=12$	−	$\frac{1}{3}[9+13+11]=11$	−

a. إذا كانت $k=2$: $MSE=\frac{2}{7}[0+2.25+0+6.25+9+0+9]=3.786$

إذا كانت $k=3$: $MSE=\frac{1}{3}[4+0+4+5+5.429+0+4]=2.905$

b. ويمكن القول بأن قيمة k = 3 أفضــل من قيمة k = 2 في التنبؤ لأن متوســط مربعات الأخطاء المناظر أقل ومن ثم فإن: $\hat{y}_{10}=\frac{1}{3}[y_9+y_8+y_7]=\frac{1}{3}[9+11+13]=11$