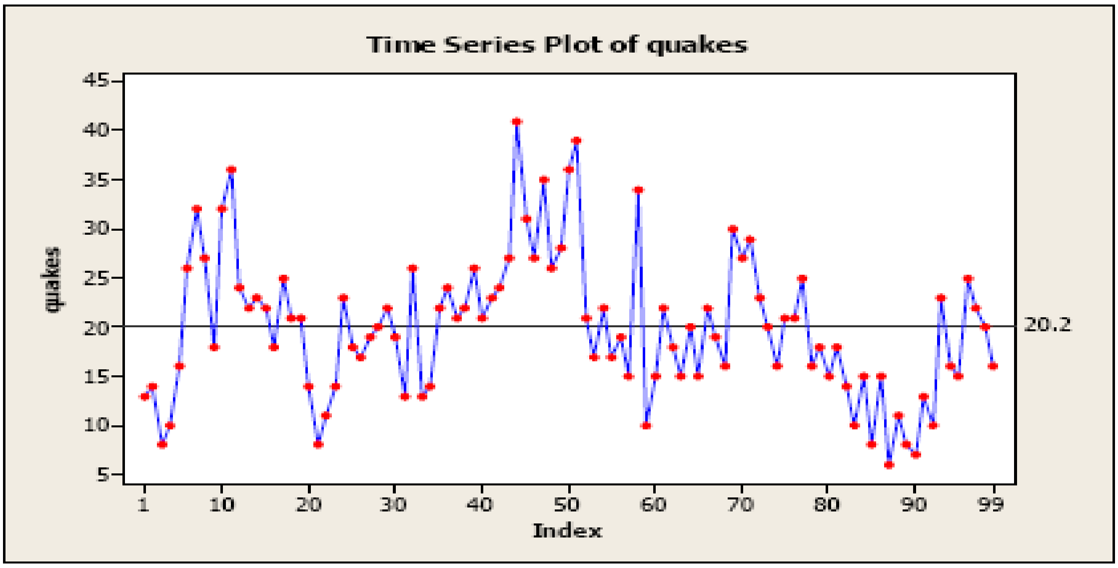
مقرر السلاسل الزمنية



المحاضرة (الأولى + الثانية)

لطلاب السنة الثالثة إحصاء رياضي

مدرس المقرر

د . رفيف الحبيب

حالات من الانحدار غير الخطي التي يمكن ردها إلى خطية:

1 الانحدار التناظري:

العلاقة التي تربط بين x و y هي:

$$(1) \ldots . \quad y = \frac{1}{a+bx} \quad ; \quad a, b \in R \quad ; \quad b \neq 0$$

لخطية هذه الدالة نضع $y^* = \frac{1}{y}$ ،وبالتالي:

$$y^* = a + bx$$

وهي معادلة خطية نعين فيها b و a بطريقة المربعات الصغرى ومن ثم التعويض عكساً في العلاقة(1) عن b و a بما يساويها للحصول على العلاقة ، وبالتالي نكون قد حصلنا على أفضل خط يمر بالبيانات المعطاة وفقاً للمربعات الصغرى.

وهذه المعادلة تساعدنا على التنبؤ بقيم y إذا كانت x معلومة ، لذلك نكتبها بالشكل:

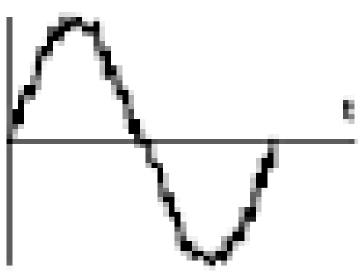
$$\hat{y} = \frac{1}{a+bx}$$

و أما من أجل تعيين قيم x إذا كانت قيم y معطاة، في هذه الحالة نستخدم العلاقة:

$$\hat{x} = \frac{1-ay}{by}$$

نحصل على هذه العلاقة باستخلاص x من العلاقة (1).

- شكل من أشكال الانحدار التناظري:



2 الانحدار اللوغاريتمي:

للمنحني الممثل لأفضل خط يمر بالبيانات العرض الآتي:

$$(2) \ldots \quad y = a + b \ln x \quad ; \quad x > 0$$

في هذه الحالة نفرض أن: $x^{*} = \ln x$

فتصبح لدينا المعادلة بالشكل: $y = a + b\, x^{*}$ وهي معادلة خطية تحسب فيها b و a بطريقة المربعات الصغرى و من ثم التعويض عكساً عن b و a في العلاقة(2) فنحصل على معادلة أفضل خط يمر من هذه البيانات وفقاً لطريقة المربعات الصغرى .

وهذه الطريقة تفيدنا في التنبؤ عن قيم y إذا كانت قيم x معلومة ، وأما إذا كانت قيم y معلومة فيمكن التنبؤ عن قيم x المقابلة وفق العلاقة:

$$\hat{x} = \exp\left[\frac{y-a}{b}\right] \quad ; \quad b \neq 0$$

نحصل على هذه العلاقة باستخلاص x من العلاقة (2).

- شكل من أشكال الانحدار اللوغاريتمي:

3 الانحدار الأسي:

في هذه الحالة سيكون لمعادلة أفضل خط يمر بهذه البيانات الشكل الآتي:

$$(3) \ldots\ y = a \,.\, e^{bx} \quad ; \quad a, b \in R^*$$

ولخطية هذه العلاقة نأخذ لوغاريتم الطرفين:

$$\ln y = \ln a + b\, x$$

ومن ثم نفرض أن: $\ln y = y^*$, $\ln a = a^*$

فنحصل على العلاقة:

$$y^* = a^* + b\, x$$

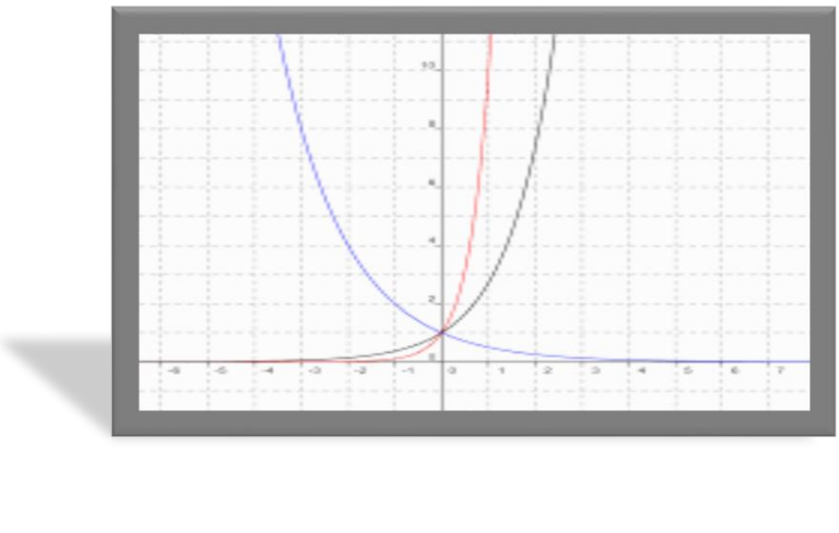
وهي علاقة خطية يحسب منها a^* و b بطريقة المربعات الصغرى ومن ثم التعويض عكساً في العلاقة (3) ، فنحصل على:

$$\hat{y} = a \,.\, e^{bx}$$

والتي تمثل معادلة أفضل خط يمر بالبيانات التي لدينا في هذه الحالة ووفقاً لطريقة المربعات الصغرى. والعلاقة (3) تساعدنا بالتنبؤ بقيم y إذا كانت x معلومة ، وأما لحساب قيم x إذا كانت قيم y معلومة ، فنستخدم العلاقة التالية:

$$\hat{x} = \frac{\ln y - \ln a}{b} \quad ; \quad a > 0\,, b \neq 0\,, y > 0$$

- شكل من أشكال الانحدار الأسي:



4 الانحدار القوي:

في هذه الحالة يكون الخط المولد للبيانات بالشكل:

$$(4) \ldots .y = a.x^b \quad ; \quad a \neq 0 , b \neq 0 , b \in N$$

ولخطية هذه العلاقة نأخذ لوغاريتم الطرفين

$$\ln y = \ln a + b \ln x$$

وبفرض أن: $\ln a = a^* \quad , \quad \ln y = y^* \quad , \ln x = x^*$

وبالتالي:

$$y^* = a^* + b\, x^*$$

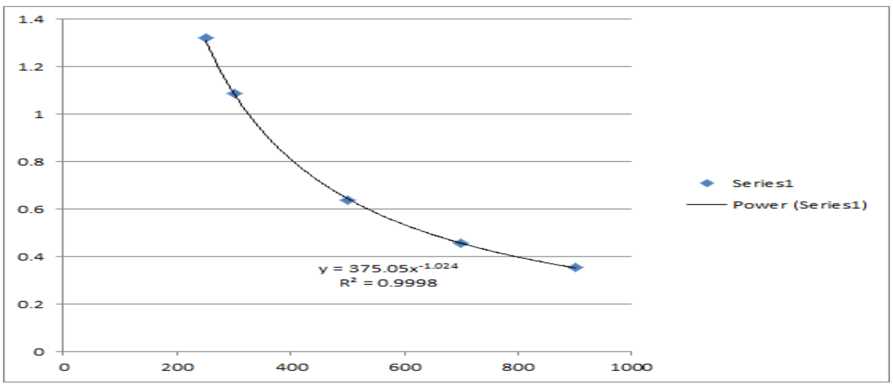
نحسب فيها a^* و b بطريقة المربعات الصغرى ومن ثم العودة عكساً للتعويض في العلاقة (4) ، فنحصل على العلاقة:

$$\hat{y} = a.x^b$$

والتي تمثل معادلة أفضل خط يمر بالبيانات المعطاة في حالتنا هذه وفق طريقة المربعات الصغرى، وهذه العلاقة تساعدنا على التنبؤ بقيم y إذا كانت x معلومة، وأما للتنبؤ عن قيم x بدلالة y فنستخدم العلاقة التالية:

$$\hat{x} = \exp\left[\frac{ln\, y - ln\, a}{b}\right] \qquad ; b \neq 0 \ , a > 0 \ , y > 0$$

- شكل من أشكال الانحدار القوي:



ملاحظة:

من نماذج الانحدار التي لا يمكن ردها إلى خطية كثيرات الحدود والدوال المثلثية.

الارتباط المتعدد والارتباط الجزئي:

✦ الارتباط المتعدد:

لقد قمنا فيما سبق بدراسة الارتباط بين ظاهرتين x و y فقط وبحثنا في تعيين قيمة العلاقة الارتباطية فيما بينها من خلال ما يعرف بمعامل الارتباط الخطي ρ أو $\rho(x,y)$، ولكن قد يصادفنا في كثير من الحالات وجود ظواهر متعددة في تجربة ما منها على سبيل المثال.. حالة غاز موجود في حيز محدود فهنا يمكن إدخال الظواهر التالية على سبيل المثال لا الحصر:

حجم الغاز كأن يكون الظاهرة x .

والضغط كأن يكون الظاهرة y .

والحرارة كأن يكون الظاهرة z.

لذلك فإن علاقة معامل الارتباط الخطي السابقة تصبح عديمة الجدوى في هذه الحالة إذا أردنا دراسة العلاقة الارتباطية بإحدى هذه الظواهر تحت تأثير بقية الظواهر الأخرى فعلى سبيل المثال.. لو أردنا دراسة العلاقة الارتباطية للظاهرة x تحت تأثير الظاهرتين y و z فهذا يعني أنه لدينا أن x هي دالة لكل من y و z معاً ... $x = f(y,z)$

إن هذه العلاقة الأخيرة تعطينا ما يعرف باسم الارتباط المتعدد.

وعلى العموم فإذا كان لدينا n ظاهرة متعلقة بموضوع بحثي ما.. فعندئذ العلاقة الارتباطية لإحدى هذه الظواهر وليكن x_k مثلاً تحت تأثير بقية الظواهر سيعطينا العلاقة:

$$x_k = f(x_1, \dots\dots, x_{k-1}, x_{k+1}, \dots., x_n) \quad ; k \in N$$

والتي هي علاقة الارتباط المتعدد لهذه الظواهر.

- وفي هذا الصدد من الممكن أن يكون للدالة f (العلاقة الارتباطية المتعددة) أشكال مختلفة سنميز منها صنفين أساسيين:

1. أن تكون الدالة f خطية عندئذ يمكننا أن نكتب :

$$x_k = a_0 + a_1x_1 + \dots + a_{k-1}x_{k-1} + a_{k+1}x_{k+1} + \dots + a_nx_n \quad ; k \in N$$

وهذه العلاقة تعطينا ما يعرف باسم الارتباط المتعدد الخطي.

2. من الممكن أن يكون للدالة f شكل كثير حدود بهذه الظواهر كأن تكون من الشكل:

$$x_k = a_0+a_1x_1{}^3 + a_2x_2{}^2 + a_3x_3 + \cdots + a_{k-1}x_{k-1}{}^l + a_{k+1}x_{k+1}{}^s + \cdots + a_nx_n{}^t$$

حيث أن: $l, s, t, k \in N$

وهذه الأخيرة تعطينا ما يعرف باسم الارتباط المتعدد اللاخطي.

ملاحظات:

1. لقد بحثنا فيما سبق، بعدما عرفنا العلاقة الارتباطية بين ظاهرتين، في مسألة توفير أفضل خط يمر بالبيانات حيث كانت لدينا النقاط الممثلة للبيانات تتوزع في فضاء ثنائي البعد. أما في حالتنا هذه بما أن النقاط التي لدينا تتوزع في فضاء بعده أكبر أو يساوي 3 فعندئذ نبحث في مسألة توفير أفضل سطح يمر بالنقاط. وفي الحالة الخاصة: عندما تكون علاقة الارتباط المتعدد خطية فعندئذ نبحث في مسألة توفير أفضل مستوي يمر بالنقاط.
2. بفرض أنه لدينا ثلاثة ظواهر فقط x_1, x_2, x_3 فعندئذ سنرمز لقيمة الارتباط المتعدد لـ x_1 تحت تأثير x_2 و x_3 بالرمز $\rho_{1,23}$ وكرمز أفضل $\rho_{1.23}$.

✦ الارتباط الجزئي:

يُعرّف الارتباط الجزئي بين متغيرات عديدة لظواهر $x_1, \dots, x_n$ بأنه العلاقة الخطية المتبادلة بين <u>متغيرين فقط</u> تحت ثبات تأثير بقية المتغيرات ، حيث تدعى هذه المتغيرات المثبت تأثيرها

بالمتغيرات المستبعدة أو المشوشة ، هذا من جانب أما من جانب آخر فإن عدد المتغيرات المستبعدة والتي في دراستنا هذه هي (n-2) يدعى بدرجة الارتباط الجزئي.

ملاحظات:

1. نرمز للارتباط الجزئي على سبيل المثال x_2 و x_1 تحت ثبات تأثير بقية المتغيرات بالشكل $x_1x_2.x_3 \dots x_n$ ، ونرمز لقيمة معامل الارتباط الجزئي بالرمز $\rho_{12.3\,4\dots n}$.
2. إذا كان تأثير المتغيرات $x_3 \dots x_n$ على x_2 و x_1 صغير لدرجة أن الارتباط الجزئي $x_1x_2.x_3 \dots x_n$ ليس ذو أهمية مقارنة بالارتباط البسيط لـ x_1x_2 عندئذ ستكون قيمة معامل الارتباط الجزئي $\rho_{12.3\dots n}$ قريبة من قيمة معامل الارتباط البسيط $\rho_{1.2} \triangleq \rho(x_1, x_2)$.
3. إذا كانت العلاقة الارتباطية بين x_1 من جهة وبقية المتغيرات من جهة أخرى $x_2, x_3, \dots, x_n$ قوية ، وكذلك كانت العلاقة الارتباطية بين x_2 من جهة وبقية المتغيرات $x_3, x_4, \dots, x_n$ من جهة أخرى أيضاً قوية فعندئذ ستكون قيمة معامل الارتباط الجزئي $\rho_{12.3\dots n}$ قريبة من الصفر.
4. إن استبعاد تأثير المتغيرات المشوشة $x_3, \dots, x_n$ سيؤثر على العلاقة الارتباطية بين x_2 و x_1 ومن ثم يكون الارتباط الجزئي $x_1x_2.x_3 \dots x_n$ سيبدو كأنه الارتباط البسيط لـ x_2 و x_1 وهذه الحالة تحدث بشكل خاص عندما تكون أحد المتغيرات x_3 أو x_4 أو .... في صيغة الارتباط المتعدد له قيمة سالبة.

5. إن القيمة العملية لمعامل الارتباط الجزئي $x_1x_2.x_3 \ \dots \ x_n$ تكون على العموم أكبر من قيمة معامل الارتباط البسيط لـ x_2 و x_1

$$\rho_{12} \leq \rho_{12.3 \ \dots \ n}$$

6. إن قيمة معامل الارتباط الجزئي تتمتع بخواص مماثلة لمعامل الارتباط البسيط ρ_{12} والتي من أهمها أن :

$$-1 \leq \rho_{12.3 \ \dots \ n} \leq +1$$

7. يجب على المرء ألا يرفع درجة الارتباط الجزئي كثيراً (أي لا يستبعد الكثير من المتغيرات المشوشة(المستبعدة)) وذلك لأن استبعاد عدد أكبر من المتغيرات المشوشة سيؤدي إلى ضعف الحكم على الارتباط الجزئي الذي هو قيد الدرس.

السلاسل (المتسلسلات) الزمنية:

تعريف: إن السلسلة الزمنية هي مجموعة من البيانات المرتبة زمنياً حيث تتميز بيانات هذه السلاسل بارتباطها ببعضها البعض في الحالة العامة وهذا الارتباط يعطينا تنبؤات مستقبلية موثوق بها. وتعرّف أيضاً بأنها مجموعة من القيم (المشاهدات) المتتالية التي تصف تطور ظاهرة ما مع الزمن.

هدف دراسة السلاسل الزمنية:

1. رصد التغيرات التي رافقت الظاهرة خلال فترة الدراسة (فترة محددة من الزمن) ووصفها وتحليلها وتصنيفها.
2. دراسة الأسباب التي أدت إلى حدوث هذه التغيرات في الظاهرة ومحاولة تقييمها بطرق علمية دقيقة.
3. التنبؤ بما سيحث من تطورات على السلسلة في المستقبل بناء على التاريخ السابق للسلسلة واعتماداً على قوانين إحصائية ورياضية تصف سلوك الظاهرة بشكل جيد في الماضي ولها القدرة على تقييم وتقدير قيمها في المستقبل بأقل قدر ممكن من الأخطاء.

أنواع السلاسل الزمنية حسب وحدة الزمن التي تقاس بها الظاهرة:

1. سلاسل زمنية تتألف من قيم مسجلة كل عشر سنوات مثل تعداد السكان وتدعى السلاسل العقدية.
2. سلاسل زمنية أخذت قيمها كل سنة مثل إنتاج القطن والقمح وتدعى سلاسل سنوية.
3. سلاسل زمنية تسجل قيمها كل فصل مثل أعداد السياح أو إنتاج بعض المحاصيل الموسمية تدعى سلاسل فصلية.
4. سلاسل زمنية تسجل قيمها بشكل شهري مثل إنتاج المعامل الشهري من الأدوية تسمى سلاسل شهرية.
5. سلاسل زمنية يومية تسجل قيمها بشكل يومي مثل درجات الحرارة ودرجات الرطوبة وسرعة الرياح وغيرها.