استخدام الانحدار الرتيب لتقدير متوسط الانفاق الشهري للأسرة العراقية لسنة 2005

Use of monotonous regression to estimate the average monthly expenditure of the Iraqi family for the year 2005

أ.م. عائدة هادي صالح
Assistant teacher aida Hadi Saleh's return
كلية الادارة والاقتصاد/الجامعةالمستنصرية
College of Administration and Economics/Al-Mustansiriya University
aida.stat@yahoo.com

الملخص

جرى الحصول على بيانات هذا البحث من المجموعة الاحصائية السنوية الصادرة من الجهاز المركزي للاحصاء والتي تضم متوسط إنفاق الاسرة الشهري على مجاميع السلع والخدمات ، ومتوسط دخل الاسرة الشهري لسنة 2005 ، لـ (18) فئة من فئات دخل الاسرة الشهري .

استُخدم في هذا البحث طريقة لامعلمية تسمى (Monotonic Regression) ، وهي طريقة سهلة ومرنة للتعامل مع البيانات الرتيبة ، حيث تستخدم البيانات المتزايدة (Increasing Data) وتصنف الى قطاعات استناداً الى سلوك البيانات ، ومن كل قطاع نحصل على خط مستقيم بمعلماته الخاصة به ، بعد ذلك نحصل على القيم المقدرة لمتغير الاستجابة من هذه المعلمات داخل كل قطاع .

استخدم في هذا البحث (Isotonic estimator) ، و اختبار (Daniel) لاختبار معنوية الاتجاه ، وكذلك تنفيذ اسلوب التصفيات الخلفية (Backward Elimination Procedure) ، لتوضيح فيما إذا كان هناك فرق بين (Isotonic Regression) و (Reduced Isotonic Regression) ، هذا الاسلوب يخفض عدد مجموعات المستويات بدمج المجموعات التي أقيامها لاتختلف كثيراً ، فقطاعات الحلول الناتجة يمكن أن تدمج سوية مع تلك التي تكون أقيامها المقدرة لاتختلف كثيراً عن القطاعات المجاورة ، أما قطاعات الحلول التي لم تبرهن بمعنوية فيمكن أن تقصى ، فالانموذج بقطاعات حلول قليلة يكون ملائم أكثر . وبعد دمج القطاعات نستخدم احصاءة الاختبار (A_m) ونقارنها مع قيمة $F(1,n-k,\alpha)$ لتحديد العدد المناسب من مجموعات المستويات ، وعند الوصول الى قيمة (A_m) أكبر من قيمة F ، نتوقف عن عملية دمج القطاعات ، وبذلك نكون قد وصلنا الى (Isotonic Regression) يختلف عن (Reduced Isotonic Regression) ، وكذلك عدد القطاعات قد انخفض ، حيث كلما انخفض عدد القطاعات كان الانموذج أفضل للتطبيق. وكذلك أوجدنا تقدير تباين متغير الاستجابة (Y) .

والهدف من هذا البحث هو استخدام طريقة احصائية لامعلمية لتحليل الانحدار اللامعلمي في تقدير متوسط الانفاق الشهري للأسرة العراقية لعام 2005 ، وذلك في حالة مشكلة عدم تحقق شروط الطرائق المعلمية في الاختبار والتقدير ، لجأنا الى اسلوب أكثر مرونة وسهولة في التعامل مع البيانات تسمى بالطرائق اللامعلمية ومنها اسلوب الانحدار الرتيب المستخدم في هذا البحث .

والاستنتاج الذي توصل اليه الباحث هو أنه يمكن تقدير متوسط الانفاق الشهري باستخدام طرائق لامعلمية بسبب المرونة العالية التي تتمتع بها المقدرات اللامعلمية قياساً بالمقدرات المعلمية .

Abstract

The data for this research was obtained from the annual statistical abstract issued from central statistical organization which includes average monthly household expenditure on totals of goods and services and the average monthly household income for the year (2005) of (18) category of the monthly family income categories .

In this research, a nonparametric method called (Monotonic Regression) is used, it is a powerful and flexible method to design monotonic data. we use increasing data and classify it into blocks according to the behavior of the data , and for each block we will get a straight line with its parameters . After that we will obtain the estimation values for the response variable from these parameters inside each block.

The methods were used in this research (Isotonic Regression Estimator), (Daniel) test, these were used to test the trend, and we implement the (Backward Elimination Procedure) to see if there is a difference between isotonic regression and reduced isotonic regression. This procedure is reduced the number of level sets by combining those whose values do not differ greatly. some of the resulting solution blocks could be pooled together especially these with the estimated values do not differ a lot from their neighboring blocks , either solution blocks do not improve significantly , the fit can be eliminated , a model with fewer solution blocks fits better . After combining the blocks, through finding a statistical (A_m), and compare it with the value of $F(1, n-k, \alpha)$ to specify the appropriate number of level set, till we reach to (A_m) is grater than (F) , we will stop combining blocks , and so we arrived to isotonic regression differ from reduced isotonic regression , as well as the number of blocks has decreased , and the model could be better for applying . And finding the estimation of the response variable (Y).

The aim of this research is to use a non-parametric statistical method in estimating the average monthly expenditure of the Iraqi family for the year 2005. And in the case of the problem of the failure of meet the requirements of the parameterized method of testing and evaluation, our committees have a more flexible and manageable approach to data called non-parametric methods, including the monotonic regression in this research .

And the researcher’s conclusion is that the average monthly household expenditure can be estimated by using non-parametric methods because of the high flexibility enjoyed by non-parametric capabilities compared to the parametric capabilities by measuring the parameter capacity.

1-مقدمة :

يعدُّ تحليل الانحدار (Analysis of Regression) من أكثر الأساليب الاحصائية استعمالا في مختلف العلوم الانسانية والاقتصادية والطبية ، يهدف الى دراسة العلاقة بين متغيرين أو اكثر مع بعضها . ويضم تحليل الانحدار متغيرين او اكثر من المتغيرات التي تربطهم علاقة منطقية تحدد فيها قيم مشاهداتها وفق التجربة تدعى بالمتغيرات التوضيحية ، وتؤثر هذه المتغيرات في تحديد قيمة متغير يدعى بالمتغير المعتمد ، والتعبير عن هذه العلاقة بين المتغير المعتمد والمتغيرات التوضيحية بصيغة رياضية تسمى نموذج الانحدار ، وتحليل العلاقة بين المتغيرات وصولاً لتحقيق ثلاثة أهداف رئيسة هي الوصف (Description) ، السيطرة (Control) ، والتنبؤ (Prediction) .

ولعل من أهم الأهداف التي تستخدم من أجلها معادلة الانحدار هي التنبؤ ، ولغرض الحصول على المعادلة التنبؤية ينبغي أولاً تقدير معاملات الانحدار في إنموذج الانحدار المفروض .

وللحصول على تقديرات جيدة لهذه المعاملات يمكن استخدام طرائق التقدير المعروفة المعتمدة على معرفة التوزيع اللاحتمالي لحد الخطأ العشوائي ، ومن هذه الطرائق طريقة الامكان الأعظم (Maximum Likelihood Method) وطريقة العزوم (Moments Method) وغيرها والتي تسمى بالطرائق المعلمية في التقدير (Parametric Methods) .

وفي الجانب العملي قد تواجهنا الكثير من المعوقات في الحصول على مثل هذه التقديرات ، فقد يكون التوزيع الاحتمالي غير معلوم ، أو أن يكون حجم العينة صغيراً ، أو أن لايكون اهتمامنا منصباً حول معالم المجتمع الذي سحبت منه العينة ، أو أن البيانات مصنفة ، أو أن المتغيرات مقاسة بالمقياس الاسمي (Nominal Scale) ، أو بالمقياس الرتبوي (Ordinal Scale) ، وبصورة عامة في حالة عدم تحقق شروط الطرائق المعلمية في التقدير والاختبار نلجأ الى اسلوب آخر قد يكون أكثر مرونة وسهولة في التعامل مع الحجوم المختلفة للعينات ، هذه الطرائق تسمى بالطرائق اللامعلمية (Nonparametric Methods) .

2-مشكلة البحث :

عدم تحقق شروط الطرائق المعلمية في التقدير والاختبار ، أو عندما يكون التوزيع الاحتمالي غير معلوم ، أو حجم العينة يكون صغيراً ، أو لايكون اهتمامنا منصباً حول معالم المجتمع الذي سحبت منه العينة ، أو أن البيانات مصنفة بسبب كون المتغيرات مقاسة بالمقياس الاسمي أو الرتبوي ، فسوف نذهب الى طريقة سهلة ومرنة في التعامل مع احجام عينات مختلفة ، وهذه الطريقة تسمى بالطريقة اللامعلمية .

3-هدف البحث :

يهدف البحث الى استخدام طريقة احصائية لامعلمية لتحليل الانحدار اللامعلمي ، عندما تكون البيانات لها صفة الرتابة (أي أن البيانات عندما يكون لها أحد اتجاهين أما بالزيادة أو النقصان) ، فطريقة التقدير المناسبة هنا هي الانحدار الرتيب (Monotonic Regression) لبيانات هذا البحث التي تتصف بالرتابة نحو الزيادة ، و استُخدم طريقة الانحدار الرتيب المتزايد (Isotonic Regression) وهي طريقة سهلة ومرنة في التعامل مع البيانات الرتيبة .

4- **الدراسات السابقة**

4-1 خلفية تاريخية حول الانحدار الرتيب History of Monotone Regression

في عام **(1955)** قدم (Brunk)[10] بحثاً في تقديرات الامكان الأعظم لمعلمات الانحدار الرتيب.

في عام **(1972)** ألف (Barlow)[9] كتاب (الاستدلال الاحصائي تحت قيود الطلب : نظرية وتطبيق الانحدار الرتيب المتزايد).

في عام **(1972)** قدما (Barlow & Brunk)[8] بحثا في مشكلة الانحدار الرتيب المتزايد المزدوج.

في عام **(1984)** قدما (Friedman & Tibshirani)[15] بحثاً أشاروا فيه الى الانحدار الرتيب والربط بين تنعيم المخططات المبعثرة والانحدار الرتيب المتزايد وكذلك قدموا فكرة حول السيطرة على أية مشكلة في إتجاه البيانات المتزايدة أو المتناقصة باستخدام خوارزمية (PAVA) .

في عام **(1988)** (Robertson)[19] نشر كتاباً قدم من خلاله معلومات عامة حول الانحدار الرتيب المتزايد وكذلك العلاقة بين التنعيم والانحدار الرتيب المتزايد.

في عام (**1988**) (Mukerjee)[18] قدم بحثاً هو دراسة إجراء هجين ينتج مقدرات رتيبة ذات خصائص مشابهة لتلك الخصائص لمقدرات الانحدار اللامعلمي أي (قام بتعديل دالة الانحدار الرتيب وحصل على تقدير رتيب بخواص مشابهة الى خواص مقدرات الانحدار الرتيب).

في عام (**1997**) قدم الباحثان (Michael & Singh)[17] بحثاً تمت فيه المقارنة بين الانحدار الرتيب المتزايد المخفض والانحدار الرتيب المخفض ، وكانت طريقة الانحدار الرتيب المخفض هي المتفوقة وطريقة مرنة لنمذجة البيانات الرتيبة .

2-4 بعض من البحوث التي نشرت بعد عام 2000

استخدم في هذا البحث بعض من البحوث كمصادر ، ونذكر فيماياْتي البعض الآخر في عام (**2005**) ، قام الباحثون (Hussain,Grimvall,Burdakov)[16] بتطوير خوارزمية (PAV) ، وكيف استطاعت أن تسهل بشكل كبير تقييم الاتجاه في السلاسل الزمنية لبيانات الجودة البيئية . وقدموا طرق حديثة للاستخراج المتزامن للاتجاه الرتيب والمكونات الموسمية ولجودة التطبيع لبيانات الجودة البيئية التي تتأثر بالتغير العشوائي في حالات الجو أو نماذج أخرى للتقلبات الطبيعية . كانت منهجية البحث مصحوبة بأمثلة لتقييمات الاتجاه لبيانات نوعية الماء وتطبيع لتركيز الزئبق في عضلة سمك القد.

في عام (**2008**) ، قدم الباحثان (دجلة ابراهيم مهدي و ياسمين عبدالرحمن محمد)[7] بحثاً جرى فيه مقارنة الطرائق اللامعلمية الرتيبة من خلال إيجاد متوسط مربعات الخطأ والكفاءة النسبية لكل مقدر ولكل إنموذج في الجانب التجريبي من خلال اسلوب المحاكاة (Monte Carlo) ، وتم مقارنة الطرائق من خلال التطبيق على بيانات عينة من (25) مريض مصابين بضغط الدم (العالي و الواطي) ، وتم التوصل الى أن طريقة (Mukerjee-stem) هي الأفضل من بين الطرائق الأخرى.

في عام (**2011**) ، قام الباحثون (Yu Han,Yunze,Yin,Xiaoming)[21] بتقييم إنجاز مقاييس جودة الصورة ، هنا وجد الانحدار الرتيب ليكون طريقة الارتباط الفعال في تقييم الأداء لمقاييس جودة الصورة . كل من التحليل النظري والنتائج التجريبية قد أثبتت بأن الانحدار الرتيب له أداء أفضل من أي انحدار آخر وأنه يمكن أن يكون أداة فعالة لتقييم الأداء في بحوث مقاييس جودة الصورة .

في عام (**2012**) ، قدم الباحثان (دجلة ابراهيم و أسماء نجم)[6] بحثاً جرى فيه تقدير الانحدار اللامعلمي باستخدام ممهد (Kernel) ، و قُدِرت الطرائق الرتيبة لجعل الدالة رتيبة متناقصة ، وجرى معالجة القيم الشاذة لجعل الدالة اللامعلمية المقدرة رتيبة باستخدام خوارزمية (PAVA) .

في عام (**2014**) ، قام الباحثان (دجلة ابراهيم و أسماء نجم)[5] بتقدير الأنواع المختلفة من حدود الثقة لمقدرات الانحدار اللامعلمية ، وجرى تقديرها باستخدام نموذج الانحدار الموضعي المتعدد الحدود والطرائق المستعملة والطرائق المقترحة عن طريق الدمج بين طريقة الانحدار الرتيب وطرائق إعادة المعاينة ، وُطبقت هذه الطرائق على بيانات حقيقية خاصة بتلوث الهواء .

5-الجانب النظري :

5-1الانحدار اللامعلمي :

هو اسلوب احصائي مغاير لمفهوم الانحدار المعلمي ويتفق معه بالهدف النهائي بالحصول على أفضل تقدير لمنحنى بيانات تتطابق أو تقترب من التطابق مع منحنى متغير الاستجابة (Y) . فلو كانت لدينا ظاهرة من الظواهر المختلفة ولاتوجد هناك أي فرضيات تحكم العلاقة بين متغير الاستجابة (Y) والمتغير التوضيحي (X_i) ، ولا يمكن تحديد أي علاقة سواء كانت علاقة خطية أو علاقة لاخطية بينهما ، عندها ستكون العلاقة المقترحة بين (Y) و (X) تسمى بالانحدار اللامعلمي

(Nonparametric Regression) ، ونقطة الخلاف الرئيسة تكون في أن الانحدار اللامعلمي لايتقيد بشروط صارمة من حيث التوزيع الخاص بمتغير الاستجابة والخطأ ، ولا يتقيد بدالة معينة تفسر العلاقة بين المتغيرات التوضيحية ومتغير الاستجابة . إنَّ المرونة العالية التي تتمتع بها المقدرات اللامعلمية قياساً بالمقدرات المعلمية ، وتطور أجهزة الحاسبات أدى الى ميل الباحثين في العقود الأخيرة للاهتمام بموضوع الانحدار اللامعلمي وطرائق التمهيد المختلفة .

هناك أنواع محتلفة من نماذج الانحدار اللامعلمي ، وتختلف طرق التقدير من إنموذج الى آخر لكنها تعتمد في مجملها على أساليب التمهيد اللامعلمي (Nonparametric Smoothing) المختلفة ، والتمهيد اللامعلمي هو توفيق المنحنيات من دون الحاجة الى إنموذج لوصف نمط متغير الاستجابة ، ويعد المدخل الرئيس للانحدار اللامعلمي الذي يهدف بالأساس الى إيجاد أفضل منحنى ممهد يطابق أو يقترب من التطابق مع منحنى متغير الاستجابة .

2-5 الانحدار الرتيب Monotonic Regression

الانحدار الرتيب طريقة لامعلمية صممت خصيصاً للتطبيقات التي تكون فيها القيمة المتوقعة لمتغير الاستجابة متزايدة أو متناقصة في العلاقات التي تحتوي على متغير توضيحي واحد أو أكثر للمتغيرات التوضيحية $(x_1, x_2, \dots, x_n)$ ، وعموماً الطريقة الحسابية الأكثر شيوعاً واستخداماً لهذا النوع من الانحدار هي معروفة بـ (Pool-Adjacent-Violators-Algorithm) وتختصر الى (PAVA) ، وهذه الخوارزمية فعالة حسابياً عندما (n=1) لأنها تزودنا بحلول مثلى وتقلل مربعات الخطأ ، أما عندما $(n>1)$ فأن (PAVA) لها فائدة مؤكدة لتقدير الاستجابات الرتيبة للمتغيرات التوضيحية التي تختلف في بعض المستويات فقط .

وعلى أية حال فأن (Burdakov) [11] وحتى وقت قريب قام بتعميم (PAVA) من الكلي الى البيانات المرتبة جزئياً بحيث أصبحت مرنة بالتعامل مع بيانات الانحدار النموذجي التي تتضمن واحد أو أكثر من المتغيرات المستمرة ، وهذا التعميم يسهل تقدير الاستجابة الرتيبة والاتجاهات المؤقتة في البيانات النوعية . فضلاً عن ذلك فأن (PAVA) لها تطبيقات مهمة في عدة مجالات في الادارة وفي الأنظمة البيئية ، وغالباً مايكون مناسب لتقدير منحنيات الجرعة-الاستجابة في الدراسات التجريبية .

تحدث العلاقات الرتيبة في مجموعة كبيرة ومتنوعة من السياقات ، فمثلاً منحنيات الجرعة والاستجابة في التجارب الطبية غالباً ما تكون رتيبة ، وكذلك درجات الحرارة ، سرعة الرياح ، وهطول الأمطار . لذلك يستلزم وجود أساليب احصائية يمكن عن طريقها تقدير النماذج التي تزيد فيها الاستجابة المتوقعة أو تنقص فيما يتعلق بواحد أو أكثر من المتغيرات التفسيرية . فالانحدار الرتيب هو طريقة لامعلمية لتقدير النماذج ، فالحالة الخاصة عندما يشار الى الاستجابة المتوقعة تتزايد في جميع المتغيرات التفسيرية فأن الانحدار يكون هو الانحدار الرتيب المتزايد Isotonic Regression حيث يمكن صياغة تقدير نموذج المشكلة التي يجري فيها تصغير دالة الخسارة ضمن مجموعة بسيطة من القيود واستجابة واحدة ولعدد من المتغيرات التوضيحية .

3-5 الانحدار الرتيب المتزايد – الحالة البسيطة المرتبة[19] Isotonic Regression-simple ordered Case

نفرض أن (X) هي مجموعة محددة $\{x_1, x_2, \dots, x_n\}$ بالترتيب البسيط $\{x_1 \le x_2 \le \dots \le x_n\}$.

الدالة f على x رتيبة ، إذا $f(x_1) \le f(x_2) \le \dots \le f(x_n)$

W يفترض أن يكون وزن دالة موجب معرف على x .

المقدر الرتيب Y^* يخفض حالة المربعات الصغرى الموزونة

$$Min\sum_{i=1}^{n}W_i\left[y(x)-f(x)\right]^2..................(1)$$

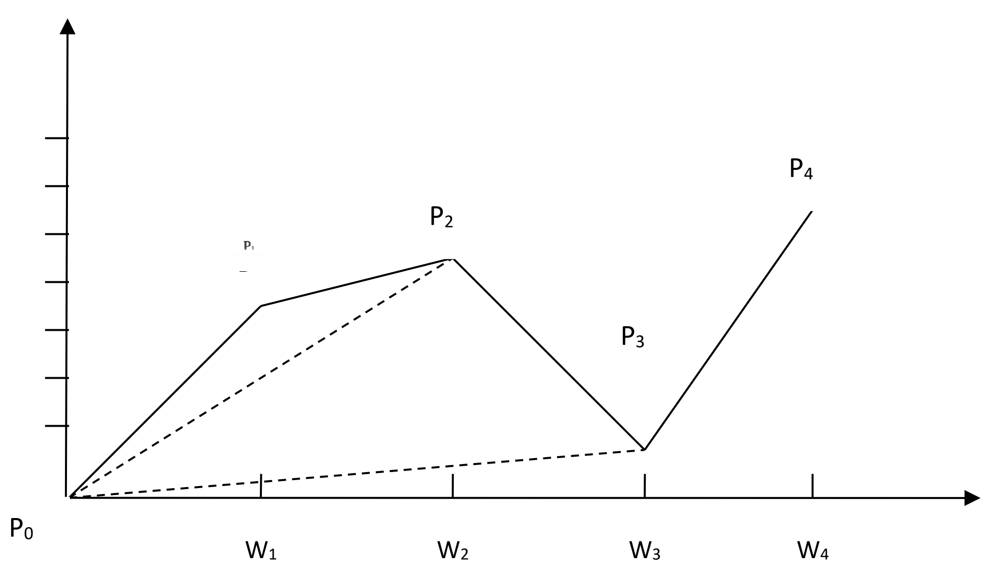
5-4 الطريقة الحسابية للخوارزمية [9],[19],[20] Computational Algorithms (Pooled Adjacent Violator Algorithm)(PAVA)

الخوارزمية البسيطة للإحتساب نوقشت بالتفصيل من قبل (Barlow) ، حيث قدم تفسير بالرسم للإنحدار الرتيب بدلالة أعظم محدب (Greatest Convex Minorant) (GCM) ، وطريقة التقريبات المتعاقبة لـــــ (GCM) يمكن أن توصـف جبرياً بـ (PAVA) ، وكذلك ذكرت في (Robertson 1988) .

(PAVA) هي خوارزمية مترابطة بسيطة لحل مشاكل الانحدار الرتيب ، والانحدار الرتيب هو مشكلة ملائمة البيانات الى قيود مرتبة ، هذه المشكلة يمكن أن تحل عددياً بإسلوب كفوء بتوقعات متتالية الى قيود بسيطة مرتبة ، هذه الخوارزمية وظفت بشكل متكرر وعلى نطاق واسع من التطبيقات .

فكرة هذا العمل هو اسلوب (Greatest Convex Minorant) واختصاراً (GCM) المستخدم هنا هو التحليل المحدب . لبعض النقاط من i ، $Y(x_{i-1})>Y(x_i)$ ، فان رسم (GCM) لـــ (W_{i-2},W_i) هو قطعة مستقيم . إذا ربطنا نقطتين P_i و P_{i-2} بقطعة مستقيم ، الـ (GCM) لهذا الرسم الجديد هو نفسه الـ (GCM) للرسم الأصلي . الـ (GCM) يمكن تكوينه بعمليات متسلسلة ، قطعة المستقيم الواحدة المجاورة في الخطوة السابقة تستبدل بقطعة مستقيم واحدة إذا كان هناك مخالف .

لنتأمل الرسم في شكل (1) الميل لقطعة المستقيم التي تربط النقطتين P_0 و P_1 أكبر من ميل قطعة المستقيم التي تربط P_1 و P_2 وهذا يشكل مخالفة . ومن ثمَّ فأن هاتين قطعتي المستقيمين تستبدل بقطعة مسقيم واحدة تربط P_0 و P_2 . والآن ميل قطعة المسقيم التي تربط P_0 و P_2 أكبر من ميل قطعة المستقيم التي تربط P_2 و P_3 ، وهكذا فأن هاتين قطعتي المستقيمين تستبدل بقطعة مستقيم واحدة تربط P_0 و P_3 وهكذا بالنسبة للبيانات فأن (GCM) سوف يحدد وبناءاً على ذلك هكذا هو الانحدار الرتيب. شكل (1) يمثل تفسير لطريقة (PAVA)



الانحدار الرتيب المتزايد (Y^*) يحدد جزء من (x) الى مجموعات من العناصـــر المتعاقبة والتي هي ثابتة . هذه المجموعات تسمى مجموعات المستويات أو قطاعات الحلول . في كل من هذه مجموعات المستويات قيمة (Y^*) هي معدل موزون لقيمة (Y) على مجموعة المستوى باستخدام الأوزان (W) ، أو بتعبير آخر قيمة (Y^*) للنقاط في مجموعة المستوى هي الميل لجانب (GCM) .

الـ (PAVA) تبدأ بـ (Y) ، إذا (Y) رتيبة (الـ CSD تكون محدبة) فأن ($Y^*=Y$) . وغير ذلك يجب إيجاد اشتراكات (i) مثل $Y(x_{i-1})>Y(x_i)$ ، هاتين القيمتين بعد ذلك تستبدلان بالمعدل الموزون لهما ، أي

$$Av(i-1,i)=\frac{[y(x_{i-1})w(x_{i-1})+y(x_i)w(x_i)]}{[w(x_{i-1})+w(x_i)]}\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

الأوزان $w(x_{i-1})$ ، $w(x_i)$ يستبدلان بـ $w(x_{i-1})+w(x_i)$ إذا المجموعة الجديدة من القيم (n−1) رتيبة بعد ذلك

$$\left.\begin{array}{l}Y^*(x_{i-1})=y^*(x_i)=Av(i-1,i)\\ Y^*(x_i)=Y(x_i)\ldots\ldots\ldots\ldots Otherwise\end{array}\right\}\quad\ldots\ldots\ldots\ldots\ldots(3)$$

إذا هذه المجموعة الجديدة من القيم غير رتيبة ، بعد ذلك تعاد العملية مســـتخدمين قيم جديدة وأوزان الى أن يجري الحصول على مجموعات من القيم الرتيبة . قيمة $Y^*(x_i)$ هي المعدل الموزون على القطاع الذي يشتمل على (x_i) . وبنفس الاسلوب يمكن الحصول على الاتجاه المتناقص .

5-5 الانحدار الرتيب المتزايد[(14),(17),(19)] Isotonic Regression

حصــل الانحدار الرتيب على اهتمام كبير جداً مؤخراً ، وأصـبح شـائع الاسـتعمال في التطبيقات الحديثة ، وله فائدة كبيرة في الاختبارات مثل اختبار الاتجاه وفي التصاميم والنماذج البديلة وفي التقديرات .

الانحدار الرتيب (Monotonic Regression) بالأســـاس يتضـــمن حالتين الأولى تســـمى الانحدار الرتيب المتزايد (Isotonic Regression)، والثانية تســـمى الانحدار الرتيب المتناقص (Antitonic Regression) ، فإذا كانت دالة الانحدار تتألف من حالة واحدة غير متناقصـــة أو غير متزايدة ، فأن طريقة التقدير المناســـبة هي فقط الانحدار الرتيب ، فطريقة الانحدار اللامعلمي هذه هي المفضـــلة على الأغلب . الانحدار الرتيب هو طريقة لامعلمية تســـتخدم عندما تكون العلاقة بين متغير الاســـتجابة المعتمد ومتغير التنبؤ المســـتقل رتيبة . تقدير الانحدار هو دالة خطوة أو(مرحلة) التي تقلل مواصـــفات (n) من النقاط الى L من مجموعات المســـتوى بحيث $(L\le n)$. هذه الطريقة تنتج نموذج يتألف من L مجتمعات فرعية متجانسة كثيرة أو قليلة . التقدير لكل مجموعة يساوي معدل متغيرات الاستجابة للنقاط في المجموعة .

إجراءات الانحدار الرتيب المتزايد هي :

$$\hat{\mu}(x_i)=a_j+b_jx_i\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

حيث إنّ :

$$a_j=\frac{\bar{y}^+(x_{m_{j-1}})+\bar{y}^+(x_{m_{j-1}+1})}{2}-\frac{1}{2}b_j\ldots\ldots\ldots\ldots(5)$$

$$b_j = \left(\frac{\bar{y}^+(x_{m_j+1}) - \bar{y}^+(x_{m_{j-1}})}{2} \right) \frac{1}{m_j} \qquad\qquad (6)$$

حيث $\hat{\mu}(x_i)$ تمثل دالة (القيم المقدرة لـ y بدلالة x) وتمثل دالة الانفاق في بحثنا هذا .

أما x فيمثل المغير التوضيحي وهو يمثل الدخل في بحثنا هذا .

أما a_j و b_j فهي معلمات .

نحصل على قيم (a_j) و (b_j) بواسطة القيم المقدرة لـ (y) بدلالة (x) و (x) هي النقاط المارة بقطع الخطوط المستقيمة .

$\hat{\mu}(x_i)$ تقدر بواسطة الخط المستقيم المار خلال النقاط

$$\left[\frac{x_{m_{j-1}} + x_{m_{j-1}+1}}{2}, \frac{\bar{y}^+(x_{m_{j-1}}) + \bar{y}^+(x_{m_{j-1}+1})}{2} \right] \quad \text{and} \quad \left[\frac{x_{m_j} + x_{m_j+1}}{2}, \frac{\bar{y}^+(x_{m_j}) + \bar{y}^+(x_{m_j+1})}{2} \right]$$

6-5 طريقة تقدير المربعات الصغرى للانحدار الرتيب المتزايد [14]

The Least Square Estimate Method of Isotonic Regression

الانحدار الرتيب المتزايد يستخدم لدوال انحدار الرتيب المتزايد أو الرتيب المتناقص . تقدير المربعات الصغرى لدالة الانحدار المتزايد $\mu(x)$ نحصل عليها استناداً الى :

نفترض لدينا توزيع احتمالي لكل قيم x_i ، $i=1,2,\dots,n$ ، كذلك نفترض لكل x_i تتوفر عينة عشوائية $\left[y_j(x_i), j=1,\dots,\eta(x_i) \right]$. عدد مشاهدات y يمكن أن يكون مختلف لمختلف x_i . الانحدار الرتيب المتزايد يطابق دالة الانحدار المستحصل عليها من $\bar{y}(x_i)$ مع الأوزان

$$w(x_i) = \frac{\eta(x_i)}{\sum_{i=1}^{N} \eta(x_i)} \qquad\qquad (7)$$

نفترض أن مشاهدات (x) رتبت بالشكل $x_1 < x_2 < \dots < x_n$. إذا $\bar{y}(x_i)$ متزايدة برتابة ، فأن معدلاتهم أنفسهم هي دالة انحدار رتيب متزايد . إذا $\bar{y}(x_i)$ لم تحقق قيود الرتابة ، نقارن المعدلات بالتعاقب الى أن تتحقق قيود الرتابة .

نفترض أن المتباينة $\bar{y}(x_k) \le \bar{y}(x_{k+1})$ لم تتحقق . بعد ذلك نقدر هاتين النقطتين استناداً الى :

$$\bar{y}(x_k) = \bar{y}(x_{k+1}) = \frac{\eta(x_k)\bar{y}(x_k) + \eta(x_{k+1})\bar{y}(x_{k+1})}{\eta(x_k) + \eta(x_{k+1})} \qquad\qquad (8)$$

قبل الاستمرار في الإجراء ، يجب فحص قيود الرتابة بين $\bar{y}(x_k), \bar{y}(x_{k-1})$ و $\bar{y}(x_{k+1})$ هل تحققت بالشكل ومن ثمَّ :

$$\bar{y}(x_1) \le \dots \le \bar{y}(x_{k-1}) \le \bar{y}(x_k) = \bar{y}(x_{k+1}) \qquad\qquad (9)$$

إذا لم تكن هذه هي الحالة ، يجب إعادة تقدير قيمة y لـ (x_{k-1}) ، (x_k) و (x_{k+1})

$$\bar{y}(x_{k-1})=\bar{y}(x_k)=\bar{y}(x_{k+1})=\frac{\eta(x_{k-1})\bar{y}(x_{k-1})+\eta x_k)\bar{y}(x_k)+\eta(x_{k+1})\bar{y}(x_{k+1})}{\eta(x_{k-1})+\eta(x_k)+\eta(x_{k+1})}.......(10)$$

والآن يجب فحص فيما إذا :

$$\bar{y}(x_1)\leq.....\leq\bar{y}(x_{k-1}2)\leq\bar{y}(x_{k-1})=\bar{y}(x_k)=\bar{y}(x_{k+1})...............(11)$$

إذا لم تكن هذه هي الحالة ، يجب إعادة تقدير قيم y لــ (x_{k-2}) ، (x_{k-1}) ، (x_k) و (x_{k+1}) باستخدام المعدل الموزون للمتوسطات ، نكرر هذه العملية الى أن تتحقق قيود الرتابة لكل قيم y المطابقة الى x_i ، $i=1,2,\ldots,k+1$. بعد ذلك نكمل فحص المتباينة فيما إذا $\bar{y}(x_{k+1})\leq\bar{y}(x_{k+2})$ قد تحققت وهكذا .

يكرر هذا الإجراء الى أن التقديرات تحقق قيود الرتابة للحصول على كل قيم x_i .

هذه التقديرات هي تقديرات المربعات الصغرى $\bar{Y}^+(x)$ لـ $\mu(x)$

5-7 الانحدار الرتيب المخفض والانحدار الرتيب المتزايد المخفض [17]

Reduced Monotonic Regression and Reduced Isotonic Regression

الانحدار الرتيب المخفض هو توســـيع من جانبين لطريقة الانحدار الرتيب المتزايد المخفض . الاتجاه العام يحدد بواسطة البيانات ، وعندما يكون الاتجاه معلوم ، فالإتجاه من جانب واحد يكون أكثر قوة لتشخيص الفروقات بين المجموعات المجاورة .

وعلى افتراض (x_i, y_i, w_i) ، $i=1,2,.....,n$ ، ترمز n الى عدد الأزواج المرتبة .

حيث $x_1\leq x_2\leq.....\leq x_n$

$$Y_i=\mu_i+\varepsilon_i \qquad , \qquad \mu_1\leq\mu_2\leq.....\leq\mu_n \qquad ....(12)$$

و W_i ، $i=1,2,.....,n$ هي الأوزان ، و $\varepsilon_i \sim iid\ N(0,\sigma^2)$

وبافتراض

$$\eta(X)\equiv E(Y/X=x)...................(13)$$

$\eta(X)$ دالة غير متناقصـة لــ (x) . نفرض (μ_i^*) ترمز الى مقدر الانحدار الرتيب المتزايد لـــ (μ_i) . إحدى الاختبارات الاحصائية للاتجاه المعنوي هو مقدر الانحدار الرتيب المتزايد باختبار الفرضيات الآتي :

$$\left.\begin{aligned}H_0:\mu_1=\mu_2=.....=\mu_n\\H_1:\mu_1\leq\mu_2\leq.....\leq\mu_n\end{aligned}\right\} \qquad ....(14)$$

وباستخدام الاحصاءة

$$\gamma_I^2=\frac{\sum_{i=1}^{n}w_i(\mu_i^*-\hat{\mu})^2}{\sum_{i=1}^{n}w_i(Y_i-\hat{\mu})^2}..............(15)$$

حيث

$$\hat{\mu} = \frac{\sum_{i=1}^{n} w_i Y_i}{\sum_{i=1}^{n} w_i} \quad ....(16)$$

قيمة γ_I^2 تقع بين الصفر والواحد وهي تمثل نسبة التغير الموضح بالإنموذج . المقدر هو مقدر المربعات الصغرى لصنف دالة غير متناقصة . وبناءً على ذلك $(\gamma_I^2) > (r_p^2)$ (بافتراض ميل الانحدار غير سالب) .

8-5 اختبار دانيال [12],[13] The Daniel's Test

مقياس الارتباط الأكثر شيوعاً واستخداماً هو معامل بيرسون للارتباط (Pearson Correlation) ويرمز له (r_p) ويعرف كالآتي :

$$r_p = \frac{\sum_{i=1}^{n} X_i Y_i - n\bar{X}\bar{Y}}{\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)^{1/2}\left(\sum_{i=1}^{n} Y_i^2 - n\bar{Y}^2\right)^{1/2}} \quad ....(17)$$

تبنى دانيال استخدام معامل ارتباط سبيرمان (r_S) لاختبار الاتجاه للقياسات المقترنة ، (X) مع الزمن (t) أو الرتبة لأية قياسات مأخوذة . اختبار الفرضيات يكون :

H_0 : لا يوجد اتجاه

H_1 : يوجد اتجاه

فرضـــية العدم هي أن قيم المتغير x مســـتقلة عن قيم المتغير y ، أما الفرضـــية البديلة تقول أن توزيع البيانات (قيم x) له علاقة بالمتغير y ، أي ليست مستقلة .

اختبار الاتجاه يعتمد على معامل ارتباط سبيرمان ، فاختبار الاتجاه من جانبين يتضمن رفض فرضية العدم التي هي لايوجد اتجاه والقبول بالفرضية البديلة أي هناك اتجاه للبيانات .

يجري الاختبار عن طريق استخدام احصاءة الاختبار (معامل ارتباط سبيرمان) ، وفق الصيغة التالية :

$$r_S = 1 - \frac{6\sum_{i=1}^{n}\left[R(X_i) - R(Y_i)\right]^2}{n(n^2-1)} = 1 - \frac{6T}{n(n^2-1)} \quad ....(18)$$

عندما يكون هناك قيمة عالية للارتباط نرفض فرضـية العدم التي تقول لايوجد اتجاه ، ونقبل الفرضـية البديلة التي تقول يوجد إتجاه .

9-5 عشوائية الاخطاء [1],[4] Randomness of the Error

الارتباط الذاتي هو حالة خاصة من الارتباط إذ يقيس درجة الارتباط بين القيم المتتالية لنفس المتغير خلال فترة زمنية محددة . معامل الارتباط الذاتي البسيط (ρ) حالة خاصـة لمعامل الارتباط البسيط (r) ، والمعاملان يتشابهان من حيث إنّ كليهما لا يناسـب العلاقات غير الخطية ، وأن (ρ) مناسـباً فقط إذا كانت قيمة (e_i) في نقطة زمنية تتبع قيمتها في النقطة

الزمنية التي تســـبقها ، وهذا هو الارتباط الذاتي من الدرجة الأولى أو الرتبة الأولى (AR1) ، قيمة معامل الارتباط الذاتي تكون $|\rho| < 1$

يمكن تقدير معامل الارتباط الذاتي (ρ) من الصيغة الأتية :

$$\hat{\rho} = \frac{\sum_{i=2}^{n} (e_i * e_{i-1})}{\sum_{i=1}^{n} e_i^2} \ldots\ldots\ldots\ldots\ldots\ldots(19)$$

ولاختبار عشوائية الاخطاء اقترح الباحثان Box & Pierce عام 1970 احصاءة اختبار الفرضية :

$$H_0 : \rho_k(\hat{e}) = 0 \qquad k = 1,2,\ldots,m$$

$$H_1 : \rho_k(\hat{e}) \neq 0 \qquad \text{for some values of } k$$

اختبار الفرضية السابقة باستخدام معاملات الارتباط الذاتي لمقدرات الاخطاء لبيان مدى كفاءة الانموذج المشخص :

صيغة الاحصاءة Q-Statistic المستعملة في الاختبار هي :

$$Q_{B-P} = n\sum_{k=1}^{m} \rho_k^2(\hat{e}) \sim \chi^2_{(m-p)} \ldots\ldots\ldots(20)$$

حيث :

n : عدد المشاهدات (حجم العينة)

m : أكبر ازاحة لـ k

P : عدد المعلمات المقدرة في الانموذج

ولغرض الاختبار تقارن قيمة Q_{B-P} المحســـوبة مع قيمة $\chi^2_{(m-P)}$ الجدولية لمســـتوى دلالة (α) فإذا كانت $Q_{B-P} < \chi^2_{(m-P)}$ دل ذلك على قبول H_0 أي أن معاملات الارتباط الذاتي للاخطاء المقدرة لها توزيع عشوائي وأن الانموذج ملائم .

وطور الباحثان Box & Ljung عام 1978 الاختبار أعلاه وتوصلا الى اختبار أكثر كفاءة صيغته :

$$Q_{B-L} = n(n+2)\sum_{k=1}^{m} \rho_k^2(\hat{e}) \ldots\ldots\ldots\ldots\ldots\ldots(21)$$

وتقارن قيمة Q_{B-L} المحســوبة مع قيمة $\chi^2_{(m-P)}$ الجدولية لمســتوى دلالة (α) ، وعند قبول H_0 يدل على أن معاملات دالة الارتباط الذاتي للأخطاء المقدرة تتوزع عشوائياً ، وأن الانموذج المشخص ملائم .

كما أن هناك عدة اختبارات للكشــف عن وجود مشــكلة الارتباط الذاتي من اهمها وأكثرها اســتخداماً مايعرف باختبار دربن–واتســون (Durbin-Watson) والذي يقيس الارتباط بين كل باقي وباقي آخر لفترة من الوقت الســابق له مباشــرة ، ولكونه مناسباً لاختبار وجود مشكلة الارتباط الذاتي من الرتبة الأولى ، بموجبه توضع الفرضية الآتية :

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

وتستخدم الأخطاء العشوائية المقدرة في حساب احصاءة (D−W)

$$(D-W) = \frac{\sum_{i=2}^{n} (e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2} \quad \ldots\ldots\ldots\ldots(22)$$

نلاحظ أن البسط يبدأ بالمشاهدة الثانية نسبة لظهور البواقي المتباطئة في البسط

$$(D-W) = \frac{\sum e_i^2}{\sum e_i^2} + \frac{\sum e_{i-1}^2}{\sum e_i^2} - 2\frac{\sum e_i^2 e_{i-1}^2}{\sum e_i^2}$$

$$(D-W) = 1 + 1 - 2\hat{\rho}$$

$$(D-W) = 2 - 2\hat{\rho}$$

$$(D-W) = 2(1-\hat{\rho}) \ldots\ldots\ldots\ldots(23)$$

هذه الاحصاءة تقيس الارتباط بين كل باقي والباقي الذي يسبقه بفترة من الزمن مباشرة .

قيمة احصاءة (D−W) تكون $0 \leq (D-W) \leq 4$

If $\rho = -1 \rightarrow (D-W) = 4$

If $\rho = 0 \rightarrow (D-W) = 2$

If $\rho = 1 \rightarrow (D-W) = 0$

كلما كانت (D−W) قريبة من الصـــفر ، دل ذلك على وجود ارتباط ذاتي موجب ، في حين كلما كانت هذه القيمة قريبة من 4 ، دل ذلك على وجود ارتباط ذاتي سالب ، واخيراً إذا اقتربت تلك القيمة من 2 ، دل ذلك على إنعدام وجود ارتباط ذاتي . ولغرض اجراء الاختبار يجب أولاَ حســاب قيمة (D−W) وفق المعادلة (23) ، ومن الجداول الخاصــة باحصــاءة (D−W) نســتخرج القيمة الدنيا (d_L) والقيمة العليا (d_U) ، بدرجة حرية مســاوية الى حجم العينة (n) ، وعدد المعلمات المقدرة (k) ، ولمستوى دلالة معين (α) ، ونقارن قيمة (D−W) المحسوبة مع القيمتين الجدوليتين الدنيا والعليا .

فإذا كانت :

ارتباط ذاتي سالب $4 - d_L < D-W < 4 \rightarrow$

ارتباط ذاتي موجب $0 < D-W < d_L \rightarrow$

لا يوجد ارتباط ذاتي $2 < D-W < 4 - d_U \rightarrow$ أو $d_U < D-W < 2 \rightarrow$

قرار غير محدد $4 - d_U < D-W < 4 - d_L \rightarrow$ أو $d_L < D-W < d_U \rightarrow$

10−5 تقدير تباين (Y) لدوال الانحدار الرتيب المتزايد [14]

Estimation Variance of (Y) for Isotonic Regression Functions

في عدة حالات يكون تقدير تباين (Y) جاهز لدينا من الدراسات الاستقصائية المماثلة . وفي حالات أخرى يمكننا عمل بعض الافتراضـــــات النظرية الواقعية ، إحدى الافتراضـــــات الشـــائعة الاســـتخدام هي كل مشـــاهدات Y ، (Y_i) ،

$i=1,2,....,n$ تمتلك نفس التباين (S^2) . فإذا كنا لانمتلك أية معلومات أولية حول التباين نستطيع بسهولة أن نقدر إذا كان هناك أكثر من مشاهدة واحدة من Y لكل (X_i) .

في بعض الحالات ، نملك فقط مشــاهدة واحدة من Y لكل (X_i) . وإذا كنا لانمتلك معلومات أولية حول التباين فسـوف تكون لدينا مشكلة كبيرة .

إذا كانت دالة الانحدار تتكون من مرحلة واحدة فقط (غير متناقصـــة) أو (غير متزايدة) ، فطريقة التقدير المناســـبة هي الانحدار الرتيب المتزايد. الطريقة التقريبية لتباين المشــاهدات طورت في هذه الحالة باســتخدام الافتراض $\sigma^2(x_i)=\sigma^2$ لجميع قيم i .

لتكن $\mu(x)$ هي دالة الانحدار الرتيب المتزايد ، ولتكن (Y_i) المشـــــاهدات المولدة من هذه الدالة . نفترض $\sigma^2(Y_i)=\sigma^2<\infty$ لكل قيم i ، فأن التباين يمكن أن يقدر بـ

$$\hat{\sigma}^2(Y_i)=\frac{1}{N}\sum_{i=1}^{N}\left[y(x_i)-\mu(x_i)\right]^2..................(24)$$

حيث $\mu(x)$ تقدر بواسطة $\hat{\mu}(x)=\bar{y}^+(x)$ ، والتباين يمكن أن يقدر بدلاً عن ذلك بـ

$$\hat{\sigma}^2_{21}(Y_i)=\frac{1}{N-p}\sum_{i=1}^{N}\left[y(x_i)-\hat{\mu}(x_i)\right]^2..................(25)$$

11-5 اسلوب التصفيات الخلفية[17] Backward Elimination Procedure

لننظر في دمج مجموعتين لمستوى معين هما m و m+1 لـ $1\le m\le L-1$ على فرض

$$\mu_i^{*m}=\begin{cases}\mu_i^* \quad , \quad if \quad i\notin B_m\cup B_{m+1}\\ \dfrac{W_mU_m+W_{m+1}U_{m+1}}{W_m+W_{m+1}} \quad , \quad if \quad i\in B_m\cup B_{m+1}\end{cases}...............(26)$$

حيث $B_1,B_2,.....,B_\ell$ ترمز الى مجموعات المستوى المرتبة للمتوسطات $\mu^*=(\mu_1^*,\mu_2^*,...,\mu_n^*)$

$$W_j=\sum_{i\in B_j}w_i...............(27)$$

وكذلك

$$U_j=\mu_i^* \qquad , \qquad i\in B_j \quad ....(28)$$

فالإنموذج المناسب الجديد سوف يفضل إذا :

$$\sum_{i=1}^{n}w_i(\mu_i^*-\hat{\mu})^2-\sum_{i=1}^{n}w_i(\mu_i^{*m}-\hat{\mu})^2...............(29).$$

كان غير معنوي احصائياً ، ولجعله معنوياً نحدد توزيع المعادلة (29) .
المعادلتين (26) و (29) يمكن أن تبسطا الى

$$\sum_{i\in B_m} W_i(\mu_i^* - \hat{\mu})^2 + \sum_{i\in B_{m+1}} W_i(\mu_i^* - \hat{\mu})^2 - \sum_{i\in C} W_i(\mu_i^{*m} - \hat{\mu})^2$$

$$\begin{aligned} &= W_m(U_m - \hat{\mu})^2 + W_{m+1}(U_{m+1} - \hat{\mu})^2 - (W_m + W_{m+1})(\hat{\mu}_m - \hat{\mu})^2 \\ &= W_m U_m^2 + W_{m+1} U_{m+1}^2 - (W_m + W_{m+1})(\hat{\mu}_m)^2 \\ &= W_m(U_m - \hat{\mu}_m)^2 + W_{m+1}(U_{m+1} - \hat{\mu}_m)^2 \end{aligned}$$

$$A_m = \sum_{i\in C}^{n} w_i(\mu_i^* - \hat{\mu}_m)^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(30)$$

حيث

$$C = B_m \bigcup B_{m+1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots(31)$$

و

$$\hat{\mu}_m = \mu_i^{*m} \quad , \qquad i \in C \qquad \ldots\ldots\ldots\ldots\ldots\ldots(32)$$

متوسط كل قيم Y في C .

نستطيع أن نشتق توزيع المعادلة (30) تحت H_0 فقط للحالة $L = 2$. في هذه الحالة المعادلة (30) تقسم على σ^2 بشرط L ، الاحصاءة تكون :

$$A_m = \sum_{i\in C} w_i(\mu_i^* - \hat{\mu}_m)^2 \div \left[\sum_{i=1}^{n} w_i(Y_i - \hat{\mu})^2 - \sum_{i=1}^{n} w_i(\mu_i^* - \hat{\mu})^2 \right] \Big/ n - L \ldots\ldots\ldots\ldots(33)$$

نلاحظ ذلك تحت H_0 :

$$\sum_{i=1}^{n} w_i(Y_i - \hat{\mu})^2 - \sum_{i=1}^{n} w_i(\mu_i^* - \hat{\mu})^2 = \sum_{i=1}^{n} w_i(Y_i - \mu_i^*)^2 \sim \sigma^2 x_{n-1}^2 \ldots\ldots\ldots(34)$$

بشرط L

$$A_m = \frac{\sum_{i\in C} w_i(\mu_i^* - \hat{\mu}_m)^2 \Big/ 1}{\sum_{i=1}^{n} w_i(Y_i - \mu_i^*)^2 \Big/ n-k} \sim F(1, n-k, \alpha) \ldots\ldots\ldots(35)$$

المعادلتين (30) ، (34) مستقلة شرطياً تحت الشرط L ، بسبب لأنهما دوال لـــــــ $(U_j, j = 1,2,\ldots..L)$ و $(Y_i - U_j, i \in B_j\, j = 1,2,\ldots.., L)$ واللتان هما مستقلتان ، وهكذا تحت H_0 ، إذا $L = 2$ ، A_m لها توزيع F بدرجتي حرية 1 و n−2 تحت شرط L .

5-12 اسلوب التصفيات الخلفية (Backward Elimination Procedure) للحصول على مقدر الانحدار الرتيب المتزايد المخفض[17],[19] :

- نحصل على الانحدار الرتيب المتزايد بمجموعات مستوى L . وهذا يكون بواسطة (PAVA) .
- نقوم باحتساب $A = Min_m A_m$ ، حيث $m = 1,2,....., k-1$ ، حيث k هو عدد مستويات المجموعات .
- إذا كان $A \leq F_{1,n-k,\alpha}$ و $A = A_m$ لبعض قيم m ، ندمج المجموعتين m و m+1 للمستوى المعين باستخدام المعادلة (26).
- نعيد الخطوات الثانية و الثالثة الى أن نصل $A > F_{1,n-k,\alpha}$.

المقدر الناتج يقطع مقدر الانحدار الرتيب المتزايد المخفض . مع ملاحظة ملاءمة الانحدار الرتيب المتزايد المخفض يطابق الانحدار الرتيب المتزايد عندما تلتزم البيانات بالاتجاه تماماً .

6- الجانب التطبيقي :

أعد هذا البحث لاختبار متوسط انفاق الأسرة الشهري على مجاميع السلع والخدمات من متوسط دخل الأسرة الشهري حسب فئات دخل الأسرة الشهري ومصدر الدخل لسنة 2005 لـــ (18) فئة من فئات دخل الأسرة [2]، باستخدام الانحدار اللامعلمي إنموذج الانحدار الرتيب .

إنموذج الانحدار المستخدم هنا لهذا الغرض هو الانحدار الرتيب المتزايد لأن البيانات تمتلك خاصية التزايد ، والهدف من هذا البحث هو استخدام طريقة احصائية لامعلمية في حالة كون البيانات لاتتبع التوزيع الطبيعي ، والعلاقة بين متغير الاستجابة والمتغير التوضيحي علاقة خطية .

جدول (1) البيانات المستخدمة في الدراسة

الدخل X (بالآلاف)	الإنفاق Y (بالآلاف)	الفئات
32770	16530	1
68140	70821	2
106406	101707	3
135200	130994	4
170037	164276	5
210312	204684	6
254225	250152	7
303315	301394	8
353839	352838	9
414772	410120	10
497677	490785	11
595523	591760	12
697118	695049	13

14	809258	818211
15	980559	984845
16	1185732	1226150
17	1539742	1581385
18	4667868	3670038
المجموع	**12964266**	

6-1 تطبيق اسلوب الانحدار الرتيب المتزايد :

باستخدام المعادلة (6)

If J=1

$$b_1 = \frac{70821-16530}{2} * \frac{1}{1} = 27145.5$$

وباستخدام المعادلة (5)

$$a_1 = \frac{(0)+16530}{2} - \frac{1}{2}(27145.5) = -5307.75$$

وباستخدام المعادلة (4)

$$\hat{Y}_1 = -37761 + 54291X$$
$$\hat{Y}_1 = 16530$$

If J=2

$$b_2 = \frac{101707-16530}{2} * \frac{1}{1} = 42588.5$$
$$a_2 = \frac{16530+70821}{2} - \frac{1}{2}(42588.5) = 22381.25$$
$$\hat{Y}_2 = 64969.75$$

وبالمثل نحتسب كل القطاعات الأتية حتى نصل الى القطاع (18)

If J=18

$$b_{18} = \frac{4667868-1539742}{1} * \frac{1}{1} = 3128126$$
$$a_{18} = \frac{1539742+4667868}{2} - \frac{1}{2}(3128126) = 1539742$$
$$\hat{Y}_{18} = 4667868$$

6-2 اختبار الاتجاه بواسطة مقدر الانحدار الرتيب المتزايد :

يمكن اختبار إتجاه البيانات بواسطة الاحصاءة γ_I^2 ، وهو إحدى الاختبارات الاحصائية لاختبار معنوية الاتجاه ، مقدر الانحدار الرتيب المتزايد ضد r_p^2 وذلك باختبار الفرضية الآتية :

$$H_0 = \mu_1 = \mu_2 = \dots = \mu_n$$
$$H_1 = \mu_1 \le \mu_2 \le \dots \le \mu_n$$

باستخدام المعادلة (7) يكون :

$$w(x_i) = \frac{1}{18}$$

وعن طريق المعادلة (16) نحصل على :

$$\hat{\mu} = \frac{12964266}{18} = 720237$$

وباستخدام المعادلة (15) نحصل على :

$$\gamma_I^2 = 0.99$$

وباستخدام المعادلة (17) نحصل على :

$$r_p = 0.99$$
$$r_p^2 = 0.98$$
$$\because \quad \gamma_I^2 > \quad r_p^2$$

نرفض فرضية العدم H_0 ونقبل الفرضية البديلة H_1 وهذا يدل على أن ميل الانحدار غير سالب أو أن البيانات لها إتجاه متزايد .

6-3 اختبار دانيل للاتجاه :

نستخدم معامل ارتباط سبيرمان (r_s) في هذا الاختبار، وباستخدام المعادلة (18) يكون :

$$r_s = 0.99$$

معامل سبيرمان كبيراً جداً وهذا يدل على أن الاتجاه معنوي ، والعلاقة قوية جداً ، أي عندما يزيد الدخل يزيد الانفاق .

6-4 اختبار الارتباط الذاتي :

لاختبار الارتباط الذاتي ، تستخدم احصاءة الاختبار Q-Statistics في المعادلة (20) ، لاختبار فيما إذا كانت الأخطاء عشوائية أو لا ، فلو كانت الأخطاء عشوائية فهذا يدل على أن الانموذج مناسب

جدول (2) الارتباطات الذاتية

Lag	Autocorrelation	Box–Ljung Statistic		
		Value	d.f	Sig.[a]
6	0.013	4.445	6	0.617
7	−0.028	4.470	7	0.724
8	−0.061	4.606	8	0.799
9	−0.089	4.924	9	0.841
10	−0.114	5.506	10	0.855
11	0.136	6.451	11	0.842
12	−0.152	7.844	12	0.797
13	−0.164	9.787	13	0.711
14	−0.170	12.402	14	0.574
15	−0.170	15.854	15	0.392
16	−0.161	20.538	16	0.197

باستخدام المعادلة (20) نحصل على :

$$Q = 18(0.8735)^2 = 15.723$$

$$x^2_{(13,0.05)} = 22.36$$

$$Q < \chi^2_t$$

نقبل فرضية العدم H_0 ونرفض الفرضية البديلة H_1 ، وهذا يدل على أن الأخطاء عشوائية وأن الانموذج مناسب .

6-5 اختبار دربن-واتسن

وباستخدام المعادلة (23) يكون :

$$(D-W) = 2(1-0.9346) = 0.1308$$

$$(D-W)_{(1,18,0.05)}$$

$$d_L = 1.16 \quad , \quad d_U = 1.39$$

نلاحظ أن القيمة المحسوبة تقع خارج المنطقة المقبولة (d_L, d_U) ، أي أن احصاءة الاختبار أصغر من (d_L). وفي هذه الحالة لانستطيع رفض فرضية العدم H_0 ، وهذا معناه أن الانموذج مناسب .

6-6 تقدير تباين (Y) لدالة الانحدار الرتيب المتزايد :

لتقدير تباين (Y) ، لدينا مستويين ، المستوى الأول تحت المقدر ، والمستوى الثاني فوق المقدر.

لتقدير تباين المستوى الأول ، نستخدم المعادلة (25) عندما $\overset{-+}{Y}(x_i) = \hat{\mu}(x_i)$ نحصل على :

$$\hat{\sigma}^2_{21}(Y_i) = 7058438527$$

ولتقدير تباين المستوى الثاني ، باستخدام المعادلة (25) عندما $\hat{\mu}(x_i) = a_j + b_j x_i$

$$\hat{\sigma}^2_{22}(Y_i) = 240877280$$

6-7 تطبيق اسلوب التصفيات الخلفية :

باستخدام المعادلة (26) لدمج مجموعتين

If m=1 m+1=2

$$\hat{\mu}_m = \hat{\mu}_1 = \mu_i^{*1}$$

$$\mu_i^{*1} = \frac{U_1W_1 + U_2W_2}{W_1 + W_2} \quad , \quad U_j = \mu_i^* \quad , \quad j=1,2$$

$$W_j = \sum_{j=1}^{2} W_j \quad , \quad W_1 = \sum_{j=1}^{1} W_1 \quad , \quad W_2 = \sum_{j=1}^{1} W_2$$

$$U_1W_1 + U_2W_2 = \sum \mu_i^* W_i + \sum \mu_i^* W_i$$

$$= 16530(1/18) + 64969.75(1/18) = 4527.7638$$

$$W_1 + W_2 = (1/18) + (1/18) = 0.1111$$

$$\mu_i^{*1} = \frac{4527.7638}{0.1111} = 40753.94$$

$$\sum W_i(\mu_i^* - \hat{\mu}_1)^2 = \frac{1}{18}\left[(16530 - 40753.95)^2 + (64969.75 - 40753.95)^2\right] = 6517804018$$

$$\sum W_i(Y_i - \mu_i^*)^2 = \frac{1}{18}(1006981874) = 55943437.46$$

وباستخدام المعادلة (35)

$$A_1 = \frac{6517804018/1}{55943437.46/18 - 16} = 1.1651$$

$$F(1,2,0.025) = 38.5$$

بما أن A_1 أصغر من F

إذاً نستمر بالمقارنة حتى نصل الى A أكبر من F ثم نتوقف

If m=2 m+1=3

$$U_2W_2 + U_3W_3 = 40753.95(\frac{1}{18}) + 101307.3(\frac{1}{18}) = 7892.29$$

$$W_2 + W_3 = (1/18) + (1/18) = 0.1111$$

$$\mu_i^{*2} = \frac{7892.29}{0.1111} = 71037.71$$

$$\sum W_i(\mu_i^{*1} - \hat{\mu}_2)^2 = \frac{1}{18}\left[(40753.95 - 71037.71)^2 + (101307.3 - 71037.71)^2\right] = 101853011$$

$$A_2 = \frac{101853011/1}{55943437.46/18 - 15} = 5.46$$

$$F(1,3,0.025) = 17.4$$

بما أن A_2 أصغر من F ، نستمر بالمقارنة

If m=3 m+1=4

$$U_3W_3 + U_4W_4 = 71037.71(\frac{1}{18}) + 131992.75(\frac{1}{18}) = 11279.47$$

$$W_3 + W_4 = (1/18) + (1/18) = 0.1111$$

$$\mu_i^{*3} = \frac{11279.47}{0.1111} = 101525.38$$

$$\sum W_i(\mu_i^{*2} - \hat{\mu}_3)^2 = \frac{1}{18}\left[(71037.71 - 101525.38)^2 + (131992.75 - 101525.38)^2\right] = 103208814$$

$$A_3 = \frac{103208814/1}{55943437.46/18 - 14} = 7.38$$

$$F(1,4,0.025) = 12.2$$

بما أن A_3 أصغر من F ، نستمر بالمقارنة

If m=4 m+1=5

$$U_4W_4 + U_5W_5 = 101525.38(\frac{1}{18}) + 166057.5(\frac{1}{18}) = 14865.72$$

$$W_4 + W_5 = (1/18) + (1/18) = 0.1111$$

$$\mu_i^{*4} = \frac{14865.72}{0.1111} = 133804.86$$

$$\sum W_i(\mu_i^{*3}-\hat{\mu}_4)^2=\frac{1}{18}\left[(10152538-11677075)^2+(166057.5-11677075)^2\right]=1478669462$$

$$A_4=\frac{147866942/1}{5594343746/18-13}=13.22$$

$$F(1,5,0.025)=10$$

نتوقف الآن لأن A_4 أكبر من F

من هذه النتيجة يتضح لنا أن الانحدار الرتيب المخفض يختلف عن الانحدار الرتيب لأن عدد القطاعات انخفض من 16 قطاع الى 13 قطاع .

7- الاستنتاجات :

1- تمكنا باستخدام مقدر الانحدار الرتيب واختبار دانيل لإختبار معنوية الاتجاه ، وظهر لنا بأن البيانات لها اتجاه متزايد.

2- تمكنا من تقدير متوسط إنفاق الأسرة الشهري بواسطة اسلوب الانحدار الرتيب ، وكذلك تمكنا من تقدير تباين متغير الاستجابة (Y) متغير الانفاق الشهري.

3- جرى تنفيذ اسلوب التصفيات الخلفية (Backword Elimination) لتوضيح فيما إذا كان هناك اختلاف بين (Isotonic Regression) و(Reduced Isotonic Regression) ، حيث إنّ (Isotonic Regression) يختلف عن (Reduced Isotonic Regression) ، لأنه يخفض عدد مجموعات المستويات بدمج المجموعات التي تكون قيمها لا تختلف كثيراً عن القطاعات المجاورة ، فقطاعات الحلول الناتجة يمكن أن تجمع سوية مع تلك التي تكون قيمها المقدرة لا تختلف كثيراً عن القطاعات المتجاورة . وبالفعل فقد جرى تخفيض عدد القطاعات الموجودة في عينة البحث من 16 قطاع الى 13 قطاع . لأنه كلما قل عدد القطاعات كان النموذج أفضل في التطبيق .

4- جرت مقارنة احصاءة (A_m) مع قيمة (F) بدرجة حرية معينة، لتحديد العدد المناسب من مجموعات المستويات ، وعند الحصول على قيمة (A_m) أكبر من قيمة (F) نتوقف عن دمج القطاعات ، وبذلك نكون قد وصلنا الى انحدار رتيب مخفض عدد القطاعات فيه قد انخفض. فالأنموذج بقطاعات حلول قليلة يكون ملائم أكثر .

8- التوصيات :

1- نوصي باستخدام الانحدار الرتيب طريقة (Reduced Isotonic Regression) لأنها تقدم قناعة كافية ومؤكدة في ايجاد الحدود القصوى الصحيحة في علاقة الجرعة والاستجابة .

2- الانحدار الرتيب هو المقترح البديل عند الرغبة في إنشاء علاقة الجرعة-الاستجابة ، وفي تصنيف المتغيرات المستمرة ، وفي تقدير قيم الحدود القصوى .

9- المصادر :

1-الزوبعي ، عبيد محمود محسن ، (2010) ، " **طريقة معدلة للكشف عن الارتباطات الذاتية واختبار دقة نموذج الانحدار** " ، مجلة العلوم الاحصائية ، العدد الثالث ، 2010 .

2-المجموعة الاحصائية السنوية ، (2005) ، الجهاز المركزي للإحصاء ، وزارة التخطيط .

3- جودة ، محفوظ ، (2007) ، "**التحليل الاحصائي الأساسي**" ، دار الوائل للطباعة ، الطبعة الأولى .

4- عبدالمنعم ، ثروت ، " **الارتباط الذاتي وطرق معالجته** " 2013/10/17.

5- مهدي و عبدالله ، دجلة ابراهيم ، أســماء نجم ، (2014) ، "**تقدير حدود الثقة لنموذج الانحدار اللامعلمي الرتيب**" ، مجلة ديالى للعلوم الصرفة ، مجلد (10) ، الاصدار (1) ، كانون الثاني ، 2014 .

6- مهدي و عبدالله ، دجلة ابراهيم ، أســماء نجم ، (2012) ، " **مقارنة بين بعض طرائق الانحدار اللامعلمي الرتيب** " ، مجلة الكوت للعلوم الاقتصادية والادارية ، عدد خاص بالمؤتمر العلمي ، مجلد (1) ، ص 116-142 ، 2012 .

7- مهدي و محمد ، دجلة ابراهيم ، ياســمين عبدالرحمن ، (2008) " **تقدير دالة الانحدار اللامعلمي باســتخدام بعض الطرائق اللامعلمية الرتيبة** " ، مجلة العلوم الادارية والاقتصادية ، مجلد (14) ، الاصدار (50) ، 2008 .

8- Barlow,R.E.; Brunk,H.D. , (1972) , “ The Isotonic Regression Problem and Its Dual “ , Journal of the American Statistical Association , 67 , 140-147.

9-Barlow,R.E.; Bartholomew,D.J.; Bermner,J.M. & Brunk,H.D. , (1972) , “ Statistical Inference Under Order Restriction : The Theory and Application of Isotonic Regression “ , Wiley , New York.

10-Brunk,H.B. , (1955) , “ Maximum Likelihood Estimates of Monotone Parameters “ , The Annals of Mathematical Statistics , 26 , 607-616 .

11-Burdakov,O.; Sysoev,O.;Grimvall,A. & Sabry,M.H. , “An Algorithm for Isotonic Regression Problem “ , The 4th European Congress of Computational Methods , 2004.

12-Conover,W.J. (1999) , “Practical Nonparametric Statistics”, John Wiley & Sons, INC.

13-Daniel Pena, George C. Tiao, Ruey S. Tasy . (2011) , “A Course in Time Series Analysis” . John Wiley & Sons, INC.

14-Dahlbom U. , (1998) , "Variance Estimation Based On Knowledge of Monotonicity and Concavity Properties" . gupea.ub.gu.se-U Dahlbom-1998, Issue Date 1/8/1998

15-Friedman ,J.; Tibshirani,R. , (1984) , “ The Monotone Smoothing of Scatter Plots “ , Thechnometrics , Vol. 26 , Issue 3 , 243-250.

16-Hussian M., Grimvall A., Burdakov O.,(2005) , “(Monotonic Reression for the Detection Temporal Trends in Enviromental Quality Data)” , Match Communications in Mathematical and in computer chemistry ,54(2005) 535-550.

17-Michael J. Schell ; Singh B. , (1997) , "The Reduced Monotonic Regression Method" , Journal of American Associated , Vol. 92, No.437 , pp 128-135 .

18-Mukerjee,H. , (1988) , “ Monotone Nonparametric Regression “ , The Annals of Statistics , Vol. 16 , No. 2 , June 1988.

19- Robertson T. ; Wright FT. and Dykstra RL. , (1988) , "Order Restricted Statistical Inference" , By Wiley , New York .

20- Salanti G. , (2003) , "The Isotonic Regression Framework , Estimating and Testing Under Order Restriction", Dissertation, LMU Munchen: Faculty of Mathematics, Computer Science and Statistics.

21-Yu Han , Yunze Cai ,Yin cao , Xiaoming Xu , (2012) ,"(Monotonic Regression : A New Way for Correlating Subjective and Objective Rating in Image Quality Research)" , IEEE Transaction on Image Proccessing , Volume 21, Issue :4 , April 2012.