جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة ميسان

كلية التربية – قسم الرياضيات

التنبؤ باستخدام نموذج الانحدار الخطي المتعدد

(دراسة تطبيقية)

بحث مقدم الى مجلس كليةالتربية جامعة ميسان وهو جزء من متطلبات نيل شهادة البكالوريوس في قسم الرياضيات

اعداد الطالبان

انعام طالب مزاوي

محمد احمد عبدالحسن

بأشراف

أ.م.د.سارة عبدالحسين بندر

1445هـ

2024م

بِسْمِ اللَّهِ الرَّحْمَٰنِ الرَّحِيمِ

(فَتَعَالَى اللَّهُ الْمَلِكُ الْحَقُّ وَلَا تَعْجَلْ بِالْقُرْآنِ مِن قَبْلِ أَن يُقْضَى إِلَيْكَ وَحْيُهُ وَقُل رَّبِّ زِدْنِي عِلْمًا)

صدق الله العظيم

الاهداء

((الشهداء اكرم منا جميها))

الى الذين بذلوا الانفس من أجل هذا الوطن والدفاع عنه

الى ارواح الشهداء جميعا لاسيما الشهيدين الصدرين والنجلين والعلوية

الطاهرة بنت الهدى رضوان الله عليها

الى الوالدين الذين بلغو اقصى درجات التعب في سبيل هذا النجاح

الى الذين لم يبخلو بمعلومة ولا نصح الى الذين بحت حناجرهم في سبيل تعليمي اصحاب

الفضل الكبير والاحسان (اساتذتي المحترمين) اخص بالذكر الدكتورة والمربية الفاضلة

أ.م.د.سارة عبدالحسين

التي لم تبخل بالنصح والتوجية الصحيح اقول قولي هذا واسئل الله العلي القدير ان

يمن علينا بالامن والامان والصحة والعافية

الشكر والتقدير

من مبدء رد الجميل وعدم نكران المعروف تعجز الكلمات عندما نتوقف

امام قامة علمية وثقافية واجتماعية وصفت بالام والمربية الحنونة قبل ان

تكون احد اهم اعمدة التعليم في جامعة ميسان فقد بذلت فيها من ابداعها لتبني جيلاً

يكون قائدا للمجتمع ويصبح نموذجاً يقتدى به كل الشكر والامتنان الى جناب

الدكتورة سارة عبدالحسين

لما قدمته لنا من معلومات قيمة كما ونشكر رئاسة قسم الرياضيات المتمثل

بالدكتور احمد كريم مطشر

ومن الذي كان حريصاً على نجاح وتفوق الطلبة

وتحية حب واحترام الى رئاسة عمادة كلية التربية المحترم المتمثلة

بالدكتور براق طالب

وفي الختام اتقدم بالشكر الجزيل والامتنان الكثير الى اعضاء لجنة المناقشة

المحترمين وفي الختام السلام والتحيه والاكرام

المستخلص

- فصل اول

يتناول موضوع البحث تحليل الانحدار المتعدد

تحايل الانحدار المتعدد هو اسلوب احصائي يمكن استخدامه لتحليل العلاقه بين متغير تابع واحد والعديد من المتغيرات المستقلة ويهدف المستقلة للتنبؤ بقيمة المتغير تابع وترجيح كل قيمة متنبئ بها مما يمكن مدى مساهمتها في التنبؤ الكلي

الفرق بين الانحدار المتعدد و البسيط

تمثل الاختلاف بين الانحدار المتعدد والانحدار البيسط في ان الاخير تنبأ بالمتغير التابع من خلال الاعتماد على علاقته بالمتغير مستقل واحد على الانحدار المتعدد الذي يعتمد على أكثر من متغير مستقل كيف يمكن استخدام الانحدار المتعدد الظرهر المستقبلية ويمكن استخدامه في حالات متنوعة

- الفصل الثاني

يتناول البحث تحليل الانحدار الخطي المتعدد يعد الانحدار الخطي المتعدد من الاساليب الاحصائية المتقدمة والتي تضمن حقه الاستدلال من اجل تحسين نتائج البحث عن طريق الاستخدام الامثل للبيانات في ايجاد علاقات بينه بين الظواهر موضوع البحث

استخدمنا الانحدار الخطي المتعدد في حاله توفير الشروط

1- ان تكون العلاقة خطيه بين المتغيرات المستقلة والمتغير تابع

2- ان تكون البيانات موزعة توزيعاً صعاً للمتغيرات المستقلة والمتغير التابع

3- يجب ان تكون قيم المتغير التابع من المستوى التداريي

استخدمنا في البحث بيانات عن الانفاق الاستهلاك العائلي Y بالالف ريال خلال شهر والدخل العائلي X وعدد افراد الاسرة X_2 .

جدول المحتويات

المقدمة

(الانحدار الخطي المتعدد)

يعد الانحدار الخطي المتعدد من الاساليب الاحصائية المتقدمة والتي تضمن دقه الاستدلال من اجل تحسين تنائج البحث عن طريق الاستخدام الأمثل للبيانات في ايجاد علاقات سيبه بين الظواهر موضوع البحث

والانحدار الخطي المتعدد هو عبارة عن إيجاد معادلة رياضية تعبر عن العلاقة بين متغيرين وتستعمل لتقديم قيم سابقة ولتنبؤ قيم مستقلة وهو عبارة ايضاً عن الانحدار للمتغير التابع (Y) على العديد من المتغيرات المستقله $(X_1, X_2, \dots, X_k)$ لذا فهو يستخدم في التنبؤ بتغيرات المتغير التابع الذي يوثر فيه عده متغيرات مستقلة اي تعتمد فكرته على العلاقات الدلالية التي تستخدام ما يعرف شكل التشتت او الانتشار فيامكاننا التنبؤ بالمستوى الرقمي في فعالية رمي المطرقه على سبيل المثال الاعتماد على دراسة حالات اخر

يمكن استخدام الانحدار الخطي المتعدد في حالة توافر الشروط التاليه :

1. ان تكون العلاقة خطية بين المتغيرات المستقلة والمتغير التابع
2. ان تكون البيانات موزعة توزيعاً طبيعياً للمتغيرات المستقلة والمتغير التابع
3. يجب ان تكون قيم المتغير التابع من المستوى التدريسي على اقل

بعد الحصول على نتائج معادلة الانحدار يجب علينا ان نبين هل ان هذه المعاملات مقبولة من الناحية الاحصائية اي معنوية احصائياً مع التنويه بان المعنوية تكون لكل معامل على حدة

الفصل الاول

"الجانب النظري"

(1-1) الانحدار المتعدد (multipl Regression)

الانحدار متعدد، هو اسلوب احصائي يمكن استخدامة لتحليل العلاقة بين متغير تابع واحد والعديد من المتغيرات المستقله ويهدف تحليل الانحدار المتعدد إلى استخدام المتغيرات المنتقلة للتنبؤ بقيمة المتغير التابع وترجيح كل قيمة متنبى بها مما يعكس مدى مساهمتها في التنبؤ الكلي [1]

(1-2) خطوات تحليل الانحدار المتعدد:

يعد تحليل الانحدار المتعدد عملية صعبة ومعقدة تتضمن العديد من الخطوات الرئيسية يمكن توضيحها من خلال مثال وكالة عقارات تحاول ايجاد طريقة التصير منازل العملاء بدقة وهذه الخطوات هي

1. جمع المعلومات عن كل متغير:

اذا ارادت الوكالة بيع 100 منزل يتطلبه منها جمع معلومات عن المتغيرات المستقلة المحددة مسبقاً لكل منزل على حدة وبالطبع سعر البيع [3]

2. استكشاف العلاقة بين كل متغير مستقل والمتغير التابع:

يمكن استخدام المعلومات التي جمعت القاء نظره على العلاقة بين حجم المنزل وسعره ومتوسط دخل الحجمي والسعر والعرب من المدارس الجيده والسعر وغيره واذا لم يثبت وجود علاقه فيكون قرار استخدام المتغير من عدم حسب طبيعة المشكلة التي تحاول الوكالة معالجتها

3. اجراء الانحداد المتعدد :

تدخل الوكالة المتغيرات التي قررت تضمينها الى البرنامج الاحصائي لتبدأ العمل مع مرعات الاخطاء والانحرافات .

تساعد هذه الخطوات في التوصل الى معادلة تساعد في معرفة الى اي مدى تمثل كل عامل من العوامل استعداداً للاختلافات في سعر المنزل وسعر المتوقع لمنزل معين مع معرفة قيمة جمع المتغيرات الآخر

- ما فرق بين الانحدار المتعدد والانحدار الخطي البسيط

تمثيل الاختلاف بين الانحدار المتعدد والانحدار الخطي البسيط في ان الاخير يتنأ بالمتغير التابع من خلال الاعتماد على علاقة بمتغير مستقل واحد على الانحدار المتعدد الذي يعتمد على اكثر من متغير مستقل

- كيف يمكن استخدام الانحدار المتعدد في التنبؤ بالظواهر المستقبلية : -

يستخدم الانحدار المتعدد من خلال تحديد العوامل التي تؤثر في المتغير التابع ثم جمع المعلومات المرتبطة بهذه العوامل سوء كانت في الماضي او معلومات حالية وبعد ذلك التنبؤ بالمستقبل من خلال استخدام هذه المعلومات

- ما المجالات التي يمكن استخدام الانحدار المتعدد فيها بطرائق متنوعة

تنوع المجالات التي يمكن استخدام تحليل الانحدار المتعدد فيها ومنها مجال الاعمال ان ثقه شركات تلجأ إلى الانحدار المتعدد التحديد تاثير السوق في أصولها وتعتمد على تحليل الانحدار للتأكد من العوامل التي توثر في اسعار اصولها واسهمها .

ومن المجالات الأخرى المجال الطبي اذ يمكن التنبؤ بمدى احتمال تعافي المريض بالاعتماد على مجموعة من العوامل مثل العمر والجنس وغير من المؤثرات وايضا يمكن استخدامه في البحوث الاجتماعية للتنبؤ سلوك افراد المجتمع والتعامل مع المواقف التي يمرون بها

- متى يتم استخدام الانحدار الخطمية المتعدد

يستخدام الانحدار متعدد الحدود بشكل شائع لتحليل البينات المنحنية وهذا يحدث عندما تكون قوة المتغير المستقل اكثر من واحد في طريقة تحليل الانحدار هذه لا يكون افضل خط ملائم ابدا هو الخط المستقيم ولكن دائماً ما يكون خط المنحنى مناسباً لنقاط البيانات

ان الانحدار الخطي المتعدد ليس مجرد اسلوب واحد وانما مجموعة من الاساليب التي يمكن ، استخدامها لمعرفة العلاقة بين متغير تابع مستقر وعدد من المتغيرات المستقلة التي عادة ما تكون مستمرة .

(1-3) الغرض من استخدام التحليل :

دراسة و تحليل اثر عدة متغيرات مستقلة كمية على متغير تابع كمي

(1-4) نموذج الانحدار الخطي المتعدد :

يمكن تمثيل نموذج الانحدار الخطي المتعدد في شكل الأتي:

$$y_i = \mu_{y_i|x_{i1},x_{i2},\dots,x_{ip}} + \varepsilon_i$$

$$= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i \qquad \dots(1)$$

حيث ان :

y_i : المشاهدة للمتغير التابع على المفردة رقم i

$x_{i1}, \beta_{i_2}, \dots, x_p$: مجموعة المشاهدات للمتغيرات المسفرة أو المستقلة على المفردة (رقم i)، وعددها p متغير مستقل

$\beta_0, \beta_1, \beta_2, \dots, \beta_p$: تعبر عن معاملات الانحدار والتي تعكس آثار المتغيرات المفسرة

ε_i : خطأ عشوائي

وبفرض ان لدينا عينة عشوائية حجمها (n) مفردة يصبح لدينا (n) من المعادلات على الصورة رقم (1) ويعبر عنها كالتالي :

$$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_p x_{1p} + \varepsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \beta_p x_{2p} + \varepsilon_2 \qquad (2)$$

$$y_n = \beta_0 + \beta_1 x_{n_1} + \beta_2 x_{n_2} + \cdots + \beta_p x_{np} + \varepsilon_n$$

تحليل الانحدار الخطي المتعدد

ومن ثم يمكن وضع المعادلات الخطية (2) في صورة معادلة مصفوفية كما يلي

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{1p} \\ 1 & x_{21} & x_{22} & x_{2p} \\ & \cdots & \cdots & \cdots \\ 1 & x_{n_1} & x_{n_2} & x_{np} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_n \end{bmatrix}$$

$$y = x\beta + \varepsilon$$

(1-5) افتراضيات النموذج

1. المشاهدات التابعة $y_1, y_2, \dots, y_n$ مشاهدات مستقلة من مجتمع له توزيع طبيعي متعدد متوسطة هو $\mu_{y|\underline{x}} = \left(\mu_{y_1|\underline{x}_1}, \mu_{y_2|\underline{x}_2}, \dots, \mu_{y_n|\underline{x}_n}\right)'$ وله مصفوفة تباين
$$\Sigma_{\underline{y|x}} = \sigma^2_{\underline{y|x}} I(n)$$
2. المشاهدات المفسرة $x_{i1}, x_{i2}, \dots, x_{ip}$ ثابتة Fixed ومعطاه كما انها مستقلة احصائيا وهذا يعني ان رتبة المصفوفة $rank\ (x) = (p+1) < n$

3. الأخطاء العشوائية $[\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]$ متغيرات عشوائية لها توزيع طبيعي متعدد متوسطة صفر وتباينة $\sum_{\underline{y|x}} = \sigma^2_{\underline{y|x}} I(n)$
4. الاخطاء العشوائية مستقلة احصائيا عن المشاهدات المفسرة .

(1-6) تقدير المربعات الصغرى (OLS) لمعاملات الانحدار

تقدير المربعات الصغرى لمتجه معاملات الانحدار $\beta = [\beta_0, \beta_1, \dots, \beta_p]'$ هو الذي يجعل مجموع مربعت الأخطاء العشوائية $SSE = \sum \varepsilon_i^2 = \varepsilon' \varepsilon$ أصغر ما يمكن، ويعبر عنه بالمعادلة التالية:

$$\hat{\beta} = (x'x)^{-1} x' y$$

$$\begin{bmatrix} \widehat{\beta_0} \\ \widehat{\beta_1} \\ \cdots \\ \widehat{\beta_p} \end{bmatrix} = \begin{bmatrix} n & \sum x_1 & \cdots & \sum x_p \\ \sum x_1 & \sum e_1^2 & \cdots & \sum x_1 x_p \\ \cdots & \cdots & \cdots & \cdots \\ \sum x_p & \sum x_1 x_p & \cdots & \sum x_p^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y \\ \sum x_i y \\ \cdots \\ \sum x_p y \end{bmatrix} \quad \dots (3)$$

ويمكن استخدام برنامج Excel في إجراء العمليات الجبرية المتعلقة بالمصفوفات

الفصل الثاني

"تطبيق للجانب النظري"

تطبيق (1) :

فيما يلي بيانات عن الإنفاق الاستهلاكي العائلي y بالألف ريال خلال الشهر، و الدخل العائلي x_1 وعدد أفراد الأسرة x_2 لـ 12 أسرة.

الانفاق الاستهلاكي	الدخل العائلي	عدد افراد الاسرة
y	x_1	x_2
2.5	9.5	1
2.8	9.8	3
3.8	10.0	4
4.5	11.1	4
4.7	12.0	4
4.7	12.8	5
4.8	13.4	5
5.2	15.4	5
6.4	15.6	6
7.1	15.9	6
8.1	20.6	6
8.9	21.8	8

والمطلوب

1. كتابة الشكل الرياضي للنموذج في المجتمع.
2. حساب تقدير المربعات الصغرى المعادلة انحدار الإنفاق على الدخل وعدد أفراد الأسرة.
3. حساب معامل التحديد وعلام يدل ؟
4. تكوين جدول تحليل التباين.
5. اختبار صلاحية النموذج الخطي.
6. تفسير معاملات النموذج واختبار معنوياتها.
7. تفسير فترات الثقة 95% للمعاملات.
8. متوسط الإنفاق المتنبأ به لأسرة دخلها 20 ألف وعدد أفرادها 7 ، وما هوا حدا الثقة لهذا المتوسط $\alpha = 0.05$

حل التطبيق

1- كتابة الشكل الرياضي للنموذج في المجتمع.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$$

2- حساب تقدير المربعات الصغرى المعادلة انحدار الإنفاق على الدخل وعدد أفراد الأسرة.

$$OLS\ Estimation \quad : \hat{\beta} = (x'x)^{-1}(x'y)$$

$$(x'x) = \begin{bmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1\, x_2 \\ \sum x_2 & \sum x_1\, x_2 & \sum x_2^2 \end{bmatrix} = \begin{bmatrix} 12 & 167.9 & 57 \\ 167.9 & 2529.83 & 866.3 \\ 57 & 866.3 & 305 \end{bmatrix}$$

$$(x'x)^{-1} = \begin{bmatrix} 1.208 & -0.105 & 0.072 \\ -0.105 & 0.024 & -0.047 \\ 0.072 & -0.047 & 0.124 \end{bmatrix}$$

$$(x'y) = \begin{bmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \end{bmatrix} = \begin{bmatrix} 63.5 \\ 973.71 \\ 337.2 \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} 1.208 & -0.105 & 0.072 \\ -0.105 & 0.024 & -0.047 \\ 0.072 & -0.047 & 0.124 \end{bmatrix} \begin{bmatrix} 63.5 \\ 973.71 \\ 337.2 \end{bmatrix} = \begin{bmatrix} -1.088 \\ 0.325 \\ 0.387 \end{bmatrix}$$

ومن ثم تأخذ معادلة التنبؤ الصورة التالية

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

$$= -1.088 + 0.325 x_1 + 0.387 x_2$$

3- حساب معامل التحديد وبيان مدلولة

$$R^2 = \frac{SSR}{SSY}$$

$$SSY = y'y - CF = \sum y^2 - \frac{(\sum y)^2}{n} = 379.23 - \frac{(63.5)^2}{12}$$

$$= 379.23 - 336.02 = 43.21$$

$$SSR = \hat{\beta}' x' y - CF = (-1.088 \quad 0.325 \quad 0.387) \begin{bmatrix} 63.5 \\ 973.71 \\ 337.2 \end{bmatrix} - 336.02$$

$$= 377.45 - 336.02 = 41.43$$

$$R^2 = \frac{SSR}{SSY} = \frac{41.43}{43.21} = 0.959$$

الدخل وعدد افراد الاسرة متغيران كمتغيران مستقلان يفسر ان 95.9% من الاختلافات في الانفاق.

4- تكوين جدول تحليل التباين

$S.O.V$	SS	df	$MS = SS/df$	f^*
$Reg.$	SSR	p	MSR	MSR/MSE
$Error$	SSE	$n-p-1$	MSE	
$Total$	SSY	$n-1$		

$$SSE = SSY - SSR = 43.21 - 41.43 = 1.78$$

$$p = 2 \quad , \quad n = 12$$

جدول تحليل التباين

$S.O.V$	SS	df	MS	f^*	Sig
$Reg_{(X_1,X_2)}$	41.43	2	20.716		0.000006
$Error$	1.78	9	0.198		
$Total$	43.21	11			

5- اختبار صلاحية النموذج الخطي عند مستوى معنوية 5%.

الفرض العدم: النموذج الخطي غير منسب للتنبؤ $H_0 : \beta_1 = \beta_2 = 0$

الفرض البديل: النموذج الخطي مناسب للتنبؤ $H_0 : at\ least\ one\ of\ \ \beta_j \neq 0\ , j = 1,2$

إحصائية الاختبار:

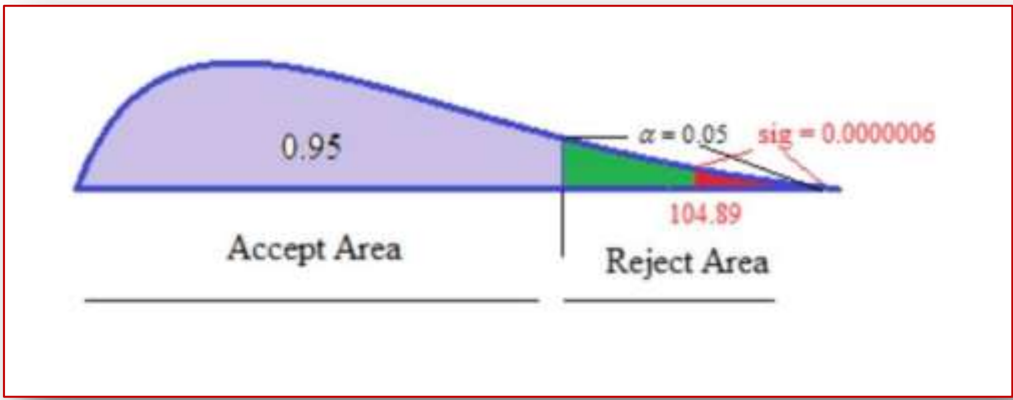
$$f^* = \frac{MSR}{MSE} = \frac{20.716}{0.198} = 104.89$$

مستوى المعنوية : $\alpha = 0.05$

الاحتمال المشاهد (المعنوية المحسوبة) :

$$p\ \ value\ (Sig.) = P\big(F_{(2,9)} > 104.89\big) = 0.0000006$$

القرار: بما أن المعنوية المحسوبة $(Sig.) = 0.0000006$ أقل من مستوى المعنوية $\alpha = 0.05$ لذا نرفض الفرض العدم ونقبل الفرض البديل، ويستدل من ذلك على أن النموذج الخطي مناسب



للتنبؤ، ويبين الشكل التالي منحنى التوزيع الاحتمالي F ومناطق الرفض والقبول.

6- تفسير معاملات النموذج واختبار معنوياتها.

لاختبار المعنويات يجب حساب مصفوفة تباين وتغاير التقديرات حتى يمكن حساب الأخطاء المعيارية لهذه التقديرات، ومصفوفة التباين والتغاير هي

$$\sum_{\hat{\beta}} = MSE(x'x)^{-1} = 0.198\begin{bmatrix} 1.208 & -0.105 & 0.072 \\ -0.105 & 0.024 & -0.047 \\ 0.072 & -0.047 & 0.124 \end{bmatrix}$$

$$= \begin{bmatrix} 0.239 & -0.021 & 0.014 \\ -0.021 & 0.005 & -0.009 \\ 0.014 & -0.009 & 0.024 \end{bmatrix}$$

وبأخذ الجذر التربيعي الموجب لكل عنصر من عناصر القطر الرئيسي نحصل على الخطأ المعياري لتقدير المعامل $S.E_{\hat{\beta}}$

$$\begin{bmatrix} S.E_{\hat{\beta}_0} & & \\ & S.E_{\hat{\beta}_1} & \\ & & S.E_{\hat{\beta}_2} \end{bmatrix} = \begin{bmatrix} \sqrt{0.239} & & \\ & \sqrt{0.005} & \\ & & \sqrt{0.024} \end{bmatrix}$$

$$= \begin{bmatrix} 0.489 & & \\ & 0.068 & \\ & & 0.157 \end{bmatrix}$$

ولاختبار معنوية المعامل ، يأخذ الفرض العدم والبديل الصورة التالية:

الفرض العدم: المتغير المستقل ليس له أثر معنوي على المتغير التابع $H_0 : \beta = 0$

الفرض البديل: المتغير المستقل ذو أثر معنوي على المتغير التابع $H_1\ :\ \beta \neq 0$

ويحسب إحصائية اختبار معنوية المعامل بالمعادلة التالية: $t^* = \hat{\beta}/S.E_{\hat{\beta}}$

والجدول التالي قيم تقديرات المعاملات ، والخطأ المعياري، وإحصائية الاختبار، وفترات الثقة 95%

للمعاملات .

					فترة ثقة 95%	
$\hat{\beta}$	$S.E_{\hat{\beta}}$	$t^* = \hat{\beta}/S.E_{\hat{\beta}}$	Sig	$t_{(9,0.975)}$	Lower	Upper
-1.088	0.489	-2.226	0.0530	2.262	-2.193	0.018
0.325	0.068	4.763	0.0010	2.262	0.170	0.479
0.387	0.157	2.472	0.0354	2.262	0.033	0.741

تفسير فترات الثقة 95% للمعاملات.

7- متوسط الإنفاق المتنبأ به الأسرة دخلها 20 ألف وعدد أفرادها 7 ، وما هوا حدا الثقة لهذا المتوسط $\alpha = 0.05$

التنبؤ بقيمة المتغير التابع Y^P عند دخل 20 الف وعدد افراد اسرة 7 x1=20, x2=7						
تقدير بنقطة Y^						
	X°=		1	20	7	
Y^^=X^° β^^			8.11243			
فترة تنبؤ للقيمة						
X° COV(β^) X(`°)		-0.076		0.006921		-0.00092
		0.05597				
	S.E_(Y^^)	0.50346				
	t_((9,0,975))	2.26216				
Lower	6.97					
Upper	9.25					
التنبؤ بقيمة متوسط المتغير التابع μP عند دخل 20 الف وعدد افراد اسرة 7 x2=7 ,x1=20						
تقدير بنقطة ^μ						
Y^=X° β^		8.11243				
فترة تنبؤ للمتوسط						
	S.E(μ^)	0.23657				
Lower	7.58					
Upper	8.65					

تطبيق (2) :

في دراسة لمتطلبات الشركات من أدوات الحاسوب اللازمة لتطوير إدارة المخزون في الشركات المتشابهة

قام الخبراء بتحديد متحولين مستقلين X_2, X_1 كما يلي:

- X_1 - عدد طلبات الزبائن (بالآلاف) .

- X_2 - عدد الطلبات المحذوفة (بالآلاف) .

كما حددوا متحولين تابعين Y_2, Y_1 كما يلي:

- Y_1 - مدة تشغيل وحدة المعالجة المركزية Cpu (بالساعات) .

- Y_2 - السعة التخزينية لقرص المدخلات والمخرجات (كيلوبايت) .

ثم قاموا بجمع البيانات اللازمة عن هذه المتحولات من (7) شركات متشابهة فكانت كما في الجدول التالي :

رقم الشركة	X_1 – الطلبات (الف)	X_2 - الطلبات المحذوفة (الف)	Y_1 - مدة تشغيل (ساعة)	Y_2 – السعة التخزينية (كيلوبايت)
1	123.8	2.108	141.5	301.8
2	146.1	9.213	168.9	396.1
3	133.9	1.905	154.8	328.2
4	128.5	0.815	146.5	307.4
5	151.5	1.061	172.8	362.4
6	136.2	8.603	160.1	369.5
7	92.0	1.125	108.5	229.1
نقطة التنبؤ x_0	130	7.5	?	?

والمطلوب إيجاد معادلات الانحدار المتعدد المضاعف للتابعين Y_2, Y_1 بدلالة المتحولين المستقلين X_2, X_1 ، ثم القيام بتقدير هذين التابعين عند النقطة $X_0' = \{1,130,7.5\}$ ثم إيجاد مجالات الثقة المتزامنة لهما . ثم إيجاد معادلة الناقص الذي يشملها باحتمال 0.95 .

الحل /

نقوم أولاً بتحديد وحساب المصفوفات اللازمة وهي:

$$X = \begin{array}{c} \begin{array}{ccc} & X_1 & X_2 \end{array} \\ \begin{bmatrix} 1 & 123.4 & 2.108 \\ 1 & 146.1 & 9.213 \\ 1 & 133.9 & 1.905 \\ 1 & 128.5 & 0.815 \\ 1 & 151.5 & 1.061 \\ 1 & 136.2 & 8.603 \\ 1 & 92.0 & 1.125 \end{bmatrix} \end{array} , Y = \begin{array}{c} \begin{array}{cc} Y(1) & Y(2) \end{array} \\ \begin{bmatrix} 141.5 & 301.8 \\ 168.9 & 396.1 \\ 154.8 & 328.2 \\ 146.5 & 307.4 \\ 172.8 & 362.4 \\ 160.1 & 369.5 \\ 108.5 & 229.1 \end{bmatrix} \end{array} = [Y(1)Y(2)]$$

$$(X'X)^{-1} = \begin{bmatrix} 8.17969 & & \\ -0.06411 & 0.00052 & \\ 0.08831 & -0.00107 & 0.0140 \end{bmatrix}$$

ثم نفترض ان معادلتي انحدار Y_1, Y_2 على X_1, X_2 كما يلي:

$$Y(1) = \beta_{01} + \beta_{11}X_1 + \beta_{21}X_2 + \varepsilon_1$$

$$Y(2) = \beta_{02} + \beta_{12}X_1 + \beta_{22}X_2 + \varepsilon_2$$

ونرمز لها مصفوفياًكما يلي:

$$Y = X * \beta + \varepsilon$$

حيث ان :

$$\beta = \begin{array}{c} \begin{array}{cc} \beta(1) & \beta(2) \end{array} \\ \begin{bmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \end{array}$$

وباستخدام الحاسوب نقوم بتقدير عناصر المصفوفة B من العلاقة (1492) فنجد أن:

$$\tilde{\beta} = (X'X)^{-1}X'Y = \begin{bmatrix} 8.42 & 14.14 \\ 1.080 & 2.25 \\ 0.42 & 5.67 \end{bmatrix}$$

وبذلك نحصل على معادلتي الانحدار المتعدد المضاعف التاليتين :

$$\tilde{Y}(1) = 8.42 + 1.080X_1 + 0.42X_2$$

$$\tilde{Y}(2) = 41.42 + 2.25X_1 + 5.67X_2$$

كما نجد أن قيمة الانحراف المعياري للأخطاء في كل منهما تساوي:

$$S_2 = 1.812 \qquad\qquad S_1 = 1.204$$

ولإيجاد القيمة التقديرية لـ $y_0(2), y_0(1)$ عند النقطة السطرية $x_0' = [1,130,7.5]$ ، نعوض ذلك في المعادلتين السابقتين، فنجد أن :

$$\tilde{y}_0(1) = 8.42 + 1.080(130) + 0.42(7.5) = 151.97$$

$$\tilde{y}_0(2) = 41.42 + 2.25(130) + 5.67(7.5) = 349.17$$

وكان يمكن حسابها من العلاقة المصفوفية التالية :

$$\tilde{y}_0 = x_0' * \beta = [1,130,7.5] * \begin{bmatrix} 8.42 & 14.14 \\ 1.080 & 2.25 \\ 0.42 & 5.67 \end{bmatrix} = \begin{bmatrix} 151.97 \\ 349.17 \end{bmatrix}$$

ولإنشاء قطع ثقة لكل من هاتين القيمتين نطبق العلاقة، لذلك نقوم أولاً بحساب الجداء

$x_0'(X'X)^{-1}x_0$ ثم حساب المصفوفة $n\tilde{V}$ فنجد أن:

$$x_0'(X'X)^{-1}x_0 = [1,130,7.5] * \begin{bmatrix} 8.17969 & -0.06411 & 0.08831 \\ -0.06411 & 0.00052 & -0.00107 \\ 0.08831 & -0.00107 & 0.01440 \end{bmatrix} * \begin{bmatrix} 1 \\ 130 \\ 7.5 \end{bmatrix}$$

$$= 0.34725$$

$$n\tilde{V} = \begin{bmatrix} \left(Y(1) - X\tilde{\beta}(1)\right)' \left(Y(1) - X\tilde{\beta}(1)\right), \left(Y(1) - X\tilde{\beta}(1)\right)' \left(Y(2) - X\tilde{\beta}(2)\right) \\ \left(Y(2) - X\tilde{\beta}(2)\right)' \left(Y(1) - X\tilde{\beta}(1)\right), \left(Y(2) - X\tilde{\beta}(2)\right)' \left(Y(2) - X\tilde{\beta}(2)\right) \end{bmatrix}$$

وبعد إجراء الحسابات اللازمة على الحاسوب نجد أن:

$$n\tilde{V} = \begin{bmatrix} 5.80 & 5.30 \\ 5.30 & 13.13 \end{bmatrix} \Rightarrow \tilde{V} = \begin{bmatrix} \frac{5.80}{7} & \frac{5.30}{7} \\ \frac{5.30}{7} & \frac{13.13}{7} \end{bmatrix}$$

ثم نقوم بتشكيل معادلة القطع الناقص للثقة للقيمة المتوقعة $E\left[Y_0\left(Y(1),Y(2)\right)\right]$ وذي الاحتمال 0.95 كما يلي:

$$[(y(1)-151.97),(y(2)-349.17)*\frac{1}{4}\begin{bmatrix}5.80 & 5.30\\ 5.30 & 13.13\end{bmatrix}^{-1}*\begin{bmatrix}Y(1)-151.97\\ Y(2)-349.17\end{bmatrix}$$
$$\leq (0.34725)\left[\frac{2(4)}{3}*F_{2\ 3}(0.05)\right]$$

وبعد تعويض $F_{2\ \ 3}(0.05)=9.55$ وإجراء بعض الإصلاحات حصلوا على قطع ناقص بدلالة $Y(2),Y(1)$، مركزه (151.97 349,17) ، ويمكن تحديد اتجاه وطول كل من محورية (الكبير والصغير) بواسطة القيم الذاتية والأشعة الذاتية للمصفوفة $\left(n\tilde{V}\right)$ كما فعلنا سابقاً

ولإنشاء مجالات الثقة المتزامنة للقيمة المتوقعة لـ $y_0(1)$ نطبق العلاقة، ولذلك نقوم أولاً بحساب ما يلي: [انظر المصفوفة $\tilde{V}$]:

$$\sqrt{x_0'(X'X)^{-1}x_0\left(\frac{n}{n-P-1}\tilde{\sigma}_{11}\right)}=\sqrt{0.34725,\frac{7}{4}\left(\frac{5.80}{7}\right)}=0.71$$

: ثم نقوم بإيجاد المقدار

$$\frac{r(n-P-1)}{n-P-n}*F_{r\ ,n-P-n}(\propto)=\frac{2(4)}{3}*F_{2\ \ 3}(0.05)=25.466$$

وهكذا نجد أن مجال الثقة ذي الاحتمال 0.95 لـ $E\left(Y_0(1)\right)$ يساوي

$$E\left(y_0(1)\right)=\tilde{Y}_0(1)\pm\sqrt{\frac{r(n-P-1)}{n-P-n}*F_{r\ ,n-P-n}(\propto)}\sqrt{x_0'(X'X)^{-1}x_0\left(\frac{n}{n-P-1}\tilde{\sigma}_{kk}\right)}$$

$$E\left(y_0(1)\right)=151.97\pm\sqrt{25.466}\sqrt{0.71}=151.97\pm 4.25$$

وهذا يعطينا المجال التالي: $146.75\leq E\left[y_0(1)\right]\leq 156.22$

وهكذا نحسب مجالات الثقة المتزامنة لـ $E[y_0(2)]$، ولإيجاد معادلة القطع الناقص التنبؤي للقيمة الفعلية لـ $Y_0\left(y_0(1),y_0(2)\right)$ لذلك نحسب أولاً قيمة المقدار :

$$1+x_0'(X'X)^{-1}x_0=1+0.34725=1.34725$$

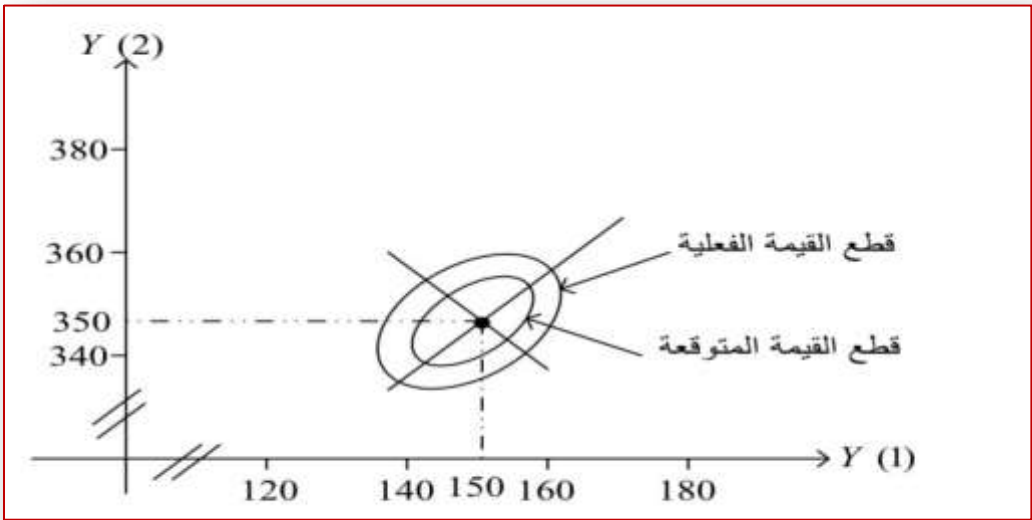
ثم نعوض ذلك في العلاقة فنجد أن :

$$[(y(1)-151.97),(y(2)-349.17)*\frac{1}{4}\begin{bmatrix}5.80 & 5.30\\5.30 & 13.13\end{bmatrix}^{-1}*\begin{bmatrix}Y_0(1)-151.97\\Y_0(2)-349.17\end{bmatrix}$$
$$\leq(1.34725)\left[\frac{2(4)}{3}*F_{2\ 3}(0.05)\right]$$

وبتعويض $F_{2\ 3}(0.05)=9.55$ وإجراء بعض الإصلاحات نحصل على قطع ناقص جديد بدلالة $y_0(1),y_0(2)$، وهو يتمركز أيضاً في النقطة (151.97,349,17) ويكون المحاوره نفس اتجاه محاور قطع الثقة السابق، ولكن أطوال محوري تكون أكبر من أطوال محوري قطع الثقة السابق . وذلك لأننا استبدلنا المقدار $x_0'(X'X)^{-1}x_0=0.34725$ بالمقدار:

$$1+x_0'(X'X)^{-1}x_0=1+0.34725=1.34725$$

وبذلك يكون القطع الثاني أكبر أو أوسع من القطع الأول. وهما يرسمان على المحورين $OY(1),OY(2)$ الشكل التالي:



ولإيجاد مجالات الثقة المتزامنة للقيمة الفعلية للتابع الأول $Y(1)$ نطبق العلاقة . لذلك نحسب أولاً المقدارين :

$$\sqrt{(1+x_0'(X'X)^{-1}x_0)\left(\frac{n}{n-P-1}\tilde{\sigma}_{11}\right)}=\sqrt{1.34725*\frac{7}{4}*\frac{5.80}{7}}=1.40$$

$$\sqrt{\frac{r(n-P-1)}{n-P-n}*F_{2\ 3}(\propto)}=\sqrt{\frac{2(4)}{3}9.55}=\sqrt{25.466}=5.05$$

وهكذا نجد أن مجال الثقة للقيمة لـ $y_0(1)$ يساوي:

$$y_0(1) = \tilde{y}_0(1) \pm \sqrt{\frac{r(n-P-1)}{n-P-n} * F_{r\,,n-P-n}(\propto)} \sqrt{(1 + x_0'(X'X)^{-1}x_0)\left(\frac{n}{n-P-1}\tilde{\sigma}_{11}\right)}$$

$$= 151.97 \pm (5.05)(1.40)$$

وهو يعطينا المجال التالي لـ $Y_0(1)$ الفعلية :

$$144.9 \leq y_0(1) \leq 159.04$$

ولإيجاد مجال الثقة المتزامن للقيمة الفعلية $y_0(2)$ نطبق نفس العلاقة لذلك نحسب المقدار :

$$\sqrt{(1 + x_0'(X'X)^{-1}x_0)\left(\frac{n}{n-P-1}\tilde{\sigma}_{22}\right)} = \sqrt{(1.34725)\left(\frac{7}{4} * \frac{13.13}{7}\right)} = 2.10$$

ولقد وجدنا سابقاً ان:

$$\sqrt{\frac{r(n-P-1)}{n-P-n} * F_2(0.05)} = 5.05$$

نعوض ذلك في العلاقة فنجد ان:

$$y_0(2) = \tilde{y}_0(2) \pm \sqrt{\frac{r(n-P-1)}{n-P-n} * F_2(0.05)} \sqrt{(1 + x_0'(X'X)^{-1}x_0)\left(\frac{n}{n-P-1}\tilde{\sigma}_{22}\right)}$$

$$= 349.17 \pm (0.05)(2.10) = 349.17 \pm 10.62$$

وهو يعطينا المجال التالي لقيمة $y_0(2)$ الفعلية:

$$338.55 \leq y_0(2) \leq 362.97$$

المراجع

1- د.ابراهيم محمد العلي ,اسس التحليل الاحصائي متعدد المتغيرات, لسنة 2020

2- د.ثائر دواد سلمان مولف الكتاب(تحليل الاحصائي للاختبارات اللامعلميه باستخدام IBM SPSS Staitstics VerSion) لعام 2020

3- د.جلال مصطفى الصياد، الإستدلال الأحصائي 2002

4- د. أسماعيل الفقهي، التحليل الإحصائي للبيانات 2012

5- د.عبد الرحمن بن محمد سليمان، الاحصاء التطبيقي 2007

6- د.محمد محمود محمد دريني ,عضو هيئه تدريس كليه العلوم عن الانحدار متعدد ,تاريخ نشر 2008