



جامعة الملك سعود

كلية العلوم

قسم الإحصاء وبحوث العمليات

دراسة تطبيقية حول اكتشاف المشاهدات القاصية والمؤثرة على الانحدار الخطي المتعدد



إعداد الطالب / عائض بن مشيب القحطاني .

بإشراف أ.د/ عبدالله بن عبد الكريم الشبيحة .

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

إهداء

الحمد لله الذي منَّ علي بإتمام بحثي هذا والذي اختص " بدراسة تطبيقية حول اكتشاف المشاهدات القاصية و المؤثرة على الانحدار الخطي المتعدد " و اشكر الله سبحانه على توفيقه و امتنانه ، لا يسعني بهذه المناسبة إلا أن أهدي ثمرة جهدي في المقام الأول لمن كان خير عون و خير مشرف و خير مرشد بل لقد كان بمثابة مرجع لي فاشكره على ما أولاني من اهتمام و حرص و على ما أمدني به من معلومات و نصائح إنه الأستاذ الدكتور / عبدا لله الشبيحة الذي أتمنى أن يتقبل مني هذا البحث تعبيراً له عن شكري و عرفاني و لا أنسى كذلك قسم الإحصاء الذي ننتمى إليه بكل فخر و اعتزاز . و كذلك لا أنسى من ساعدني و كان خير عون لي في هذا المشروع الأخ / مزهود بن سعيد الشمrani . و كذلك الأخ / يوسف بن علي القحطاني و الأخ / ناصر بن فاضل الشمري فاشكرهم جزيل الشكر .

1,II المقدمة وأهداف البحث	•
1 تحديد مشاهدات قاصية في X (المتغير المستقل) - مصفوفة القبة وقيم العزم	1
1 حالات قاصية (شاذة)	
2 استخدام مصفوفة القبة "H" وقيم العزم (القيم النفوذية أو الرافعة) h_{ii}	
3 طريقة كشف القيم القاصية في المتغيرات المستقلة باستخدام قيم العزم h_{ii} (القيم النفوذية أو الرافعة)	
33 تحديد مشاهدات قاصية في Y (المتغير التابع) - بواقي الحذف المعيرة	2
33 حساب رواسب الحذف	
34 بواقي الحذف المعيرة (بواقي ستودنت المحذوفة)	
35 طريقة كشف المشاهدات القاصية في المتغير التابع Y	
45 تحديد المشاهدات المؤثرة - تدابير (IDENTIFYING INFLUENTIAL CASES) (DFFITS , COVRATIO, Cook's Distance, DFBETAS,)	3
45 التأثير على قيم التوفيقية - مقياس (DFFITS)	3-1
46 طريقة كشف التأثير لقيمة DFFITS على النموذج	
46 التأثير على معاملات الانحدار - مقياس (DFBETAS)	3-2
47 طريقة كشف التأثير لقيمة DFBETAS على النموذج	
47 قياس الأثر على كل معاملات الانحدار - مقياس مسافة كوك	3-3
48 طريقة كشف التأثير لقيمة مسافة كوك D_i	
48 قياس الأثر على الأخطاء المعيارية - مقياس COVRATIO	3-4
49 طريقة كشف التأثير لقيمة COVRATIO على الأخطاء المعيارية	
67 تشخيصات الارتباطية الخطية المتعددة - تضخم التباين	4
68 عامل تضخم التباين (VIF)	
69 طريقة الكشف عن وجود ارتباط خطي متعدد	
69 متوسط قيم عوامل تضخم التباين لمعاملات الانحدار	
81 المراجع	

المقدمة :-

- تحليل الانحدار:

هو أداة إحصائية تستفيد من العلاقة بين متغيرين كميين أو وصفيين أو أكثر للتنبؤ بأحد المتغيرات استنادا إلى قيم المتغير أو المتغيرات الأخرى فمثلا إذا علمنا العلاقة بين مصروفات الدعاية وبين المبيعات فيمكننا الاستفادة من تحليل الانحدار للتنبؤ بالمبيعات حالما تتوفر لنا قيمة نفقات الدعاية .

لكن التحليل الإحصائي قد يتأثر ببعض المشاهدات القاصية (الشاذة) أو المتطرفة في المتغير المستقل X أو المتغير التابع Y التي قد يكون لها تأثير واضح على تحليل الانحدار .

ومن أسباب بروز المشاهدات القاصية حدوث أخطاء إما في مرحلة جمع البيانات أو في مرحلة المعالجة كإدخال البيانات في الحاسب . وفي هذا البحث سوف نتطرق إلى التشخيصات المحسنة للتحقق من صلاحية نموذج الانحدار ويتضمن البحث عن المشاهدات القاصية والمؤثرة في المتغير المستقل X أو المتغير التابع Y وكذلك سيتم دراسة مشكلة الارتباطية الخطية المتعددة وأيضا سوف ندرس كيف ستتم التدابير العلاجية لها بإذن الله .

أهداف البحث :-

- تحسينات لصلاحية نموذج الانحدار الخطي .
- طرق كشف المشاهدات القاصية (الشاذة) في المتغير المستقل X والمتغير التابع Y .
- طرق كشف المشاهدات المؤثرة على نموذج الانحدار الخطي .
- تشخيصات الارتباطية الخطية المتعددة - تضخم التباين .

البرامج المستخدمة :-

- Microsoft office Excel
- MINITAB 14
- SPSS 13
- SAS