بسم الله الرحمن الرحيم

جامعـة السودان للعلوم والتكنولوجيا

كليـة الدراســات العليــا والبحث العلمي

بحث تكميلي مقدم لنيل درجة الماجستير في الاحصاء التطبيقي بعنوان:

مقارنة بين نموذج الانحدار الخطي المتعدد ونموذج الدالـة التمييزية

" دراسـة تطبيقيـة علـي عينـة مـن الاصـحاء ومرضي- الفشـل الكلوي بمستشفي ابن سينا 2010 م "

Comporison Between The Linear Model and Discriminatory Analysis Models

Applied Study on healthy and Renal Failure Samples in "Ibn Sina Hospital 2010".

إعـــــداد:

ابراهيم محمد ابراهيم سيدأحمد

إشـــــراف:

د-عادل موسي يونس

بسم الله الرحمن الرحيم

﴿ سَنُرِيهِمْ آيَاتِنَا فِي الْآفَاقِ وَفِي أَنفُسِهِمْ حَتَّى يَتَبَيَّنَ لَهُمْ أَنَّهُ الْحَقُّ أَوَلَمْ يَكْفِ بِرَبِّكَ أَنَّهُ عَلَى كُلِّ شَيْءٍ شَهِيدٌ ﴾

صدق الله العظيم

سورة فصلت، الآية رقم (53)

إلى والدتي التي سهرت الليالي لتوفر لنا الراحة

إلى روح والدي المرحوم الذي أخذ بيدي في هذه الحياة أسال الله له جنات الفردوس باذن الله

إلى روح المرحوم أخي الذي لم تلده أمي :الخضر عبدالله مبشر لما بذله مع من جهد حتى آخر لحظات حياته

إلى أخوتي واخواتي الأعزاء الذين تحملوا العبء عني أثناء الدراسة

إلى أختي التي لم تلدها أمي:

نسيبة محمد أحمد خضر التي لم تبخل علي بشيءاحتجت اليه في دراستي

إلى أصدقائي : أشرف،محمد,عمر,مجدي ,أبوبكر,النعمه وجميع أصدقائي الذين حفظهم الله ورعاهم

إلى زملائي الأعزاء الذين سهلوا لي أموراً لم يكن من المقدرة لتجاوزها دون تعاونهم

" رَبِّ أَوْزِعْنِيَ أَنْ أَشْكُرَ نِعْمَتَكَ الَّتِيَ أَنْعَمْتَ عَلَيَّ وَعَلَىَ وَالِدَيَّ وَأَنْ أَعْمَلَ صَالِحاً تَرْضَاهُ وَأَدْخِلْنِي بِرَحْمَتِكَ فِي عِبَادِكَ الصَّالِحِينَ " وأخص بشكري الدكتور المبجل / عادل موسى يونس المشرف على البحث ، نسأل الله أن يجعل ذلك في ميزان حسناته .

و أخص بالشكر الأستاذ / محمد الأمين لتوجيهاته القيمة في بداية هذا العمل، و أخص بالشكر أيضاً الدكتور المخلص / مبارك حسن لمساعدته لي بكل ما يملك , و أخص بالشكر الأخ/أحمد عبد الله مبشر .

و أتقدم بأسمى آيات الشكر لكل من مد لي يد العون مساعداً و مساهماً برأيه حتى أكتمل هذا البحث .

الباحث

هدف هذا البحث إلى تطبيق طريقة إحصائية علمية متمثلة في النموذج الانحدار الخطي المتعدد والتحليل التمييزي المتعدد لتصنيف المصابين من غير المصابين بأمراض الكلي ومعرفة المتغيرات التي تؤثر علي الاصابة بأمراض الكلي ومن ثم المقارنة بينهما ويتم ذلك ببناء نموذج انحدار خطي متعدد ودوال تمييزية تساعد الطبيب المختص في التشخيص الأولي وتم استخدام طرق مختلفة للتمييز بين المجموعات والمقارنة بينها لتحديد الطريقة الأفضل ثم استخدام هذه النماذج التمييزية في تصنيف الحالات الجديدة إلى مجموعاتها الصحيحة.

وتم تقسيم البيانات الي مجموعتين مجموعة المصابين بأمراض الكلي ومجموعة غير المصابين بأمراض الكلي المأخوذة من مستشفي بن سينا.

وبعد ذلك تم تطبيق النموذجين علي هذه البيانات وتوصل البحث إلى نتائج مفادها:

1- أن الانحدار الخطي المتعدد له قدرة علي تمييز المصابين من غير المصابين بأمراض الكلي باستخدام القيم التقديرية .

2- أن الانحدار الخطي المتعدد أفضل من التحليل التمييزي في التمييز بالرغم من ان الدالة التمييزية التي الحصول عليها لها قدرة علي التمييز.

3- اليوريا والكرتاتنين والفوسفات لها تأثير معنوي علي الاصابة والتسبب بأمراض الكلي بينما الصوديوم ليس له تاثير معنوي علي الاصابة والتسبب بأمراض الكلي.

4- زيادة المتغيرات التي يتم قياسها لمعرفة الاصابة بهذا المرض الخطير لأن هناك متغيرات لم تدخل في النموذج حيث أن معامل التحديد التي الحصول عليه يساوي 87%, وعمل المزيد من الدراسات التي تقارن بين النماذج الاحصائية.

وبناء على نتائج البحث توصي الدراسة بالآتي:

1. استخدام نموذج الانحدار الخطي المتعدد للتمييز بين مجموعتين أو اكثر وذلك من خلال القيم التقديرية.
2. إذا تم استخدام الدالة التمييزية للتمييز بين مجموعتين أو اكثر يستخدم معها الانحدار الخطي المتعدد لمعرفة المتغيرات التي لاتؤثر في النموذج واستبعادها ومعالجة مشاكل التداخل الخطي والارتباط الذاتي بين المتغيرات المستقلة بعد ذلك يمكن ان نستخدم التحليل التمييزي باطمئنان.

3. زيادة المتغيرات (العوامل) التي تؤخذ من الاشخاص لان هناك متغيرات اخري لها أثر في المرض غير موجودة , و اجراء المزيد من الدراسات والبحوث حول الدالة التمييزية ومعرفة المشاكل التي قد تعترض لها الدالة التمييزية لمعالجتها .

4. المقارنة بين أسلوبين أو أكثر لمعرفة الاسلوب الافضل حتي تعم الفائدة للجميع مثلا اجراء دراسات ومقارنات حول الدالة التمييزية والانحدار اللوجستي والانحدار الخطي والنماذج غير الخطية

5. الفحص والتشخيص المبكر للوقاية من شر هذا المرض الخطير وعمل فحوصات واختبارات دورية لمستويات اليوريا والكرتاتنين والبوتاسيم .

The objective of this research is to apply scientific statistical method represented in the form of multiple linear regression and multiple discriminatory analysis for classification of the injured from non-infected kidney disease and find out the variables that affect the kidney disease and then the comparison between them, is by building a model of multiple linear regression & discriminatory functions to help the specialist in diagnose the first diagnosis, was the use of different methods to discriminate between groups and compare them to determine the best way and then use these models in the discriminatory classification of new cases to their properly groups.

The Data were divided into two groups, one group of kidney disease and another group of not infected with kidney disease taken from the abn- Siena Hospital.

After that the two models were applied to these data and the results of research found that:

1- The multiple linear regressions has the ability to discriminate infected from non-infected kidney disease using estimated values.

2- The best multi-linear regression analysis is better than the discriminatory function in classify, although the function of discriminatory accessed is has ability to discrimination.

3- Alkrtatnen and the urea and phosphate have significant effect on the incidence and cause of kidney disease, while sodium has no significant effect on the incidence and cause of kidney disease.

4- increase the variables that are measured to see cases of this dangerous disease because there are variables not included in the model , where the coefficient of

determination to get it equal to 87%, and work more studies that compare statistical models.

Based on the findings of the study recommends the following:

1. using multiple linear regression model to discriminate between two or more, through the estimated values.

2- if used the discriminatory function to discriminate between two or more we might use the multiple linear regression to determine variables that do not affect in the model and exclusion and addressing the problems of multicolinearty and auto correlation between independent variables, after that we can use the discriminatory analysis safely.

3- Increasing the variable (factors) taken of people because there are other variables have an effect on the disease does not exist, and do further studies and research on the function discriminatory, and know more knowledge about their problems to treatment.

4- Comparison between method or more to find the best method, so that a benefit for all, like studies and comparisons about the discriminatory function and logistic regression and linear regression models and non-linear

5- checking and early diagnosis it's good to for prevent from evil of this dangerous disease and the checking and periodic tests for levels of urea & potassium and Alkrtatnen.

الرقم	الموضوع	الصفحة
1	الآية	ب
2	الإهداء	ج
3	الشكر والتقدير	د
4	المستخلص	هـ - و
5	Abstract	ز-ح
6	الفهرست	ط-ي-ك
7	فهرست الأشكال والجداول	ل
8	**الفصل الأول**	
9	مقدمة	1
10	أهداف البحث	1
11	مشكلة البحث	1

12	أهمية البحث	2
13	فروض البحث	2
14	حدود البحث	3
15	منهج البحث	3
16	محتويات الدراسة	4
17	الدراسات السابقة	4
18	**الفصل الثاني**	
19	أمراض الكلي	
20	المقدمة	6
21	تعريف الفشل الكلوي	6
22	الفشل الكلوي الحاد	7
23	الفشل الكلوي المزمن	11
24	مسببات الفشل الكلوي	13
25	**الفصل الثالث**	
26	نموذج الانحدار الخطي المتعدد	
27	طبيعة نموذج الانحدار	17
28	نموذج الانحدار الخطي بوجود((k من المتغيرات المستقلة	17
29	فرضيات النموذج	18
30	تقدير النموذج بصيغة الانحرافات	22

22	النموذج الخطي المتعدد ذو المقدرات القياسية	31
23	تحليل التباين	32
27	اختبار تأثير كل متغير من المتغيرات المستقلة علي حده	33
28	مشكلات النماذج القياسية	34
33	التنبؤ	35
37	معامل عدم التساوي لثايل	36
	نموذج التحليل التمييزي	37
39	مقدمه	38
39	الدالة التمييزية	41
40	حساب الدالة التمييزية	42
47	الأهمية النسبية للمتغيرات	43
47	إختبار قدرة دالة التمييز في التنصيف بين المجموعتين	44
53	علاقة التحليل التمييزي بالإنحدار الخطي	45
54	نقطة الفصل	46
55	نسبة الخطأ	47
	الفصل الرابع	48
	التحليل الإحصائي	49
58	وصف متغيرات الدراسة	50
60	تحليل نموذج الانحدار الخطي المتعدد	51

52	تحليل الدالة التمييزية	71
53	المقارنة بين نموذج الانحدار الخطي المتعدد والتحليل التمييزي	82
54	التنبؤ لبعض القيم باستخدام النموذجين	84
55	**الفصل الخامس**	
56	النتائج والتوصيات	
57	الإستنتاجات	86
58	التوصيات	87
59	المصادر والمراجع	89
60	الملاحق	90

الصفحة	الموضوع	الرقم

ة		
	فهرست الجداول	1
58	يوضح وصف متغيرات مجموعة المصابين بأمراض الكلى	2
59	يوضح وصف متغيرات مجموعة غير المصابين بأمراض الكلى	3
60	يوضح معلمات النموذج	4
61	يوضح تحليل التباين	5
64	ملخص نموذج الانحدار	6
66	قيمة دربن واتسون	7
67	قيم عامل تضخم التباين	8
68	القيم التقديريه للمصابين	9
69	القيم التقديريه لغير المصابين	10
70	نتائج اختبار Kolmogorov	11
74	قيم المتوسطات للمصابين وغير المصابين	12
75	معكوس المصفوفه	13
76	أهمية متغيرات دالة التمييز	14
77	القيم التميزيه للمصابين	15
78	القيم التميزيه لغير المصابين	16
78	اختبار T	17

18	Willas' Lambd اختبار	79
19	جدول تحليل التباين	80
20	المقارنه بين الانحدار والتحليل التمييزي	81
21	بعض القيم لمتغيرات الدراسه المأخوذه من مصابين وغير مصابين بالكلى	82
22	التنبؤ	83
	فهرس الاشكال	
1	الكشف عن مشكلة عدم تجانس التباين	65
2	مدى تقارب القيم الحقيقيه مع التقديريه	68
3	رسم الاحتمال الطبيعي	69
4	شكل Box	71