SUST
Journal of Natural and Medical Sciences

Journal homepage: http://journals.sustech.edu/

استخدام التحليل التمييزي المتعدد لتصنيف مراحل الإصابة بمرض الفشل الكلوي المزمن (دراسة تطبيقية في مستشفى احمد قاسم لإمراض القلب وزراعة الكلى في الفترة من 2005م – 2012م)

عفراء هاشم عبد اللطيف و أمينة سليمان حسين*

Using of Multiple Discriminant AnalysisTo classify Chronic Renal failure(Applied Study in Ahmed Gassim to Hospital for Heart and Growing Renal)

جامعة السودان للعلوم والتكنولوجيا / كلية العلوم – قسم الإحصاء التطبيقي
aminasuliman1@hotmail.com

Article history: Recieved: 09.06.2014
Accepted: 09.07.2014

المستخلص

تم استخدام التحليل المميز المتعدد لتصنيف مراحل الإصابة بمرض الفشل الكلوي المزمن اعتماداً على متغيرات ذات صفات تمييزية. تضمنت الدراسة سحب 322 كعينة عشوائية، وباستخدام الدالة المميزة الخطية+ تم تصنيف مراحل الإصابة بمرض الفشل الكلوي المزمن على أساس مجموعة من المتغيرات المصاحبة للمرض . الغرض الأساسي لهذا العمل هو بناء دوال تمييزية تساعد الطبيب المختص في تشخيص المرض وأيضا تشخيص العوامل المؤثرة في مرض الفشل الكلوي ومعرفة تأثير كل عامل من هذه العوامل من خلال التحليل التمييزي المتعدد لتصنيف الإصابة بمرض الفشل الكلوي المزمن حسب مراحله المتعددة .

الكلمات المفتاحية : التحليل التمييزي المتعدد، الفشلالكلويالمزمن، الدالة المميزة الخطية .

ABSTRACT

The disciminant analysis has been used to classifychronic renal failure diseases based on change of disciminant feature . The study has been employed a random sample by using the linear disciminant function chronic renal failure have been classified on the base of the disease accompanied change. The assential purpose for this work is to create disciminant analysis to help a doctor to know the affective reasons in a chronic renal failure throw disciminant analysis to classified a renal failure throw it's steps.

KEYWORDS: Multiple Discriminant Analysis **,** Chronic Renal Failure**,** TheLinear Disciminant Function**.**

المقدمة

هدفت هذه الدراسة إلى التعرف على كيفية استخدام التحليل التمييزي المتعدد في تحليل مرض الفشل الكلوي نسبة لإهمية هذا المرض القاتل وانتشاره بصوره واسعة في الآونة الأخيرة .

ركزت هذه الدراسة في استخدام التحليل التمييزي المتعدد لتمييز المصابين بمرض الفشل الكلوي عن غيرهم وتمكن أطباء الكلى ومن يهمهم الأمر، لمعرفة هذا المرض من مراحله الأولية حيث يعتبر التحليل التمييزي Discriminant Analysis من أساليب التحليل التمييزي الإحصائي متعدد المتغيرات وهو من الطرق الإحصائية المهمة التي تهتم بفصل مجموعات مختلفة من المفردات أو" المشاهدات " وبتوزيع المفردات أو " المشاهدات " الجديدة على مجموعات سبق تعريفها [(1)].

تضمن هذا العمل استخدام التحليل التمييزي المتعدد لتمييز المصابين بمرض الفشل الكلوي وذلك وفقاً للمراحل المتعددة لهذا المرض ، لان الكلى لا تقوم بوظائفها كما يجب في ترشيح الدم وتخليص الجسم من المواد الضارة وذلك وفقاً للمراحل المتعددة لهذا المرض ، ولذلك يتراكم في انسجه وأعضاء الجسم المختلفة ، لذلك نهتم بتطبيق التحليل التمييزي المتعدد لمعرفة الأسباب المؤثرة على الإصابة بة ويتم ذلك ببناء دوال تمييزية تستخدم للكشف عن حالات الإصابة بهذا المرض .إن هنالك نوعان من الفشل الكلوي Renal Failure الفشل الكلوي الحاد Acute Renal Failure يظهر هذا النوع نتيجة أسباب عده قد لا تكون للكلية علاقة بها ومن الجيد أن هذه الأسباب معروفه ويمكن في كثير من الأحيان الوقاية منها ومعالجتها . والنوع الثاني الفشل الكلوي المزمن Chronic Renal Failure يعتبر الفشل الكلوي المزمن مستمر تلقائياً غير رجعي في الوحدات الوظيفية للكلية (النفرونات) ، مما يزيد من تراكم الايض (الهدم) والبناء الخلوي ، السوائل ، واليوريا في الجسم والتي تسبب مجموعه من الأعراض والعلامات ، نعرف أن مرض الفشل الكلوي من اخطر الأمراض ، علينا الابتعاد عن جميع مسببات المرض[2] .

الطرق والاساليب المستخدمة

تم تحليل البيانات التي جمعت وتم فحصها باستخدام أساليب رياضيه مستخدمه بواسطة الحاسوب واستنباط المعلومات المتوفرة فيها واتخاذ القرار.

يعمل البرنامج الإحصائي على تفسير وشرح أسباب الاختلاف في مفردات البيانات الإحصائية عبر العوامل التي ترتبط بها ، واهم برامج الحاسوب استخداما حاليا هوحزمه التحليل الإحصائي للعلوم الاجتماعية SPSSوحيث تم تحليل بيانات هذا البحث بواسطة هذا البرنامج.

وتختص هذه البيانات بمرضى الفشل الكلوي المزمن حيث تم اخذ العينة من مستشفى احمد قاسم للقلب وزراعة الكلى.

الأساليب الإحصائية المستخدمة

تم التحليل على أساس خمس مجموعات

لتحقيق هدف الدراسة وللتحقق من فرضياتها تم استخدام الأساليب الإحصائية الآتية:

1 –وصف متغيرات البحث عن طريق إيجاد جداول التكرارات والنسب
2 –إيجاد دوال التمييز ثم التصنيف
3 –إيجاد المصفوفات .
4 –تحديد المتغيرات ذات التأثير المعنوي باستخدام التمييز القويم
5 –تحديد المتغيرات ذات التأثير المعنوي باستخدام مجموع المربعات التنبؤي

وللحصول على نتائج دقيقه قدر الإمكان ، تم استخدام برنامج SPSS وبرنامج Minitab

الدالة المميزة الخطية في حالة أكثر من مجموعتين

نفرض لدينا K من المجموعات، وكل مجموعة لها n من المشاهدات، وكل مشاهدة تتضمن P من المتغيرات .

n_i : هوحجم العينة المسحوبة من المجموعة i[3].

$$n = \sum_{i=1}^{k} n_i$$

حيث S هي مصفوفة التشتت لداخل المجموعات للمتغيرات $\underline{x}$ ويكون تباين بين المجموعات للدالة الخطية هو:

$$D = \underline{\alpha}' B \underline{\alpha}$$

حيث B هي مصفوفة مصفوفة التشتت لبين المجموعات للمتغيرات $\underline{x}$ والتي عناصرها هي :

$$b_{jj} = \frac{\sum_{i=1}^{g} n_i (\overline{x}_{ij} - \overline{x}_j)^2}{g-1} , b_{jk} = \frac{\sum_{i=1}^{g} n_i (\overline{x}_{ij} - \overline{x}_j)(\overline{x}_{ik} - \overline{x}_k)}{g-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$\overline{x}_{ij} \equiv$ متوسط المجموعه i للمتغير $\overline{x}_j$

$\overline{x}_j \equiv$المتوسط العام للمتغير $\overline{x}_j$

تستخدم في التحليل المميزالاختبلوات التالية:

1-اختبارمعنوية الدالة المميزةالخطية.

2-اختبارتساوي مصفوفات التباين والتباين المشترك

$$E = \frac{\lambda_i}{\sum_{i=1}^{r} \lambda_i} 100\% \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

التصنيف في حالة عدة مجاميع

سنناقش هنامسألة تصنيف مشاهدة معينة إلى واحد من k من المجاميع التي تتوزع توزيعا طبيعي متعدد المتغيرات ، بمتجهات المتوسطات$\underline{\mu}_1, \underline{\mu}_k \ldots\ldots$,Σ_2على التوالي ومصفوفة تباين مشتركة لكل المجاميعΣ. وفي حالة كون معلمات هذه المجاميع مجهولة فأنه يمكن استخدام تقديراتها من العينة[1]:

$$\hat{\underline{\mu}}_j = \overline{\underline{X}}_j$$

$$S = \frac{1}{N-K} \sum_{j=1}^{k} A_j \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

حيثإن:

$$D_i^2 = (\overline{x} - \overline{\underline{x}})' S^{-1} (\underline{x} - \overline{\underline{x}_i}) \quad \ldots\ldots\ldots\ldots(4)$$

D_i^2 : يمثل Mahalanobis Distance

وان قاعدة التصنيف ستكون بأننصنف المشاهدة $\underline{x}$ إلى المجموعة i منخلا لاختيارأقل القيم للمقياس D_i^2مقارنة بكل المجاميع[4].

حيث C_iثابت ويستخرج كالأتي :

$$C_i = -\frac{1}{2} \overline{\underline{x}}' S^{-1} \overline{\underline{x}}_i \quad \ldots\ldots\ldots\ldots\ldots\ldots(5)$$

كما إن متجه المعاملات يستخرج كالأتي :

$$\alpha_i = \overline{\underline{x}}' S^{-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots(6)$$

$$P(\pi_i / x) = \frac{p_i e^{Z'_i}}{\sum_{i=1}^{g} p_i e^{Z'_i}} \ldots\ldots\ldots\ldots\ldots\ldots(7)$$

اختبار ويلكس لمدا (Λ) Wilks' Lambda test

هو اختبار لمعرفة نسبة الاختلافات داخل المجموعات إلى الاختلافات بين المجموعات ويعبر عنه كالاتي [4] :

$$\text{Wilks' Lambda} = \Lambda = \frac{\text{Variation Between Groups}}{\text{Variation Within Groups}} = \frac{SSb}{SSw} \quad \ldots\ldots\ldots\ldots\ldots\ldots(8)$$

وتكون قيمة الاختبار محصورة بين الصفر والواحد أيأن $(0<\Lambda>1)$.

ويستخدم في التحليل التمييزي لمعرفة معنوية الدوال التي كونت بواسطة التمييز القويم .

ففي التحليل التمييزي القويم يعتبر تقريب لاختبار $\lambda_{(p-s)(q-s)}$ – مربع كاي – وياخذ ألصيغه:

$$\Lambda = \{n - (p + q + 3)/2\} \ln r \quad \ldots\ldots\ldots\ldots(9)$$

حيث :

P:عدد متغيرات المجموعة الأولى $x_{(1)}$

q: عدد متغيرات المجموعة الثانية $x_{(2)}$

r:قيمة الارتباط القويم

وهو أيضا يدخل في حساب اختبار Rao's ، تقريب لاختبار F ، ويأخذ الصيغة .

$$\text{Rao's F}= \left[\frac{r_{2s}}{i_{1s}}\right]\left[\frac{1-\Lambda_s^{1/m_s}}{\Lambda_s^{1/m_s}}\right] \quad \ldots\ldots\ldots\ldots\ldots\ldots(10)$$

حيث :

$$m_s = \left\{\frac{(p-s)^2\,(q-s)^2-4}{(p-s)^2+(q-s)^2-5}\right\}^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(11)$$

$$r= (p-s)\,(q-s)$$

$$r_{2s}= m_s[\, n-(p+q+3)/2]-[(p-s)(q-s)-2]/2 \quad \ldots\ldots\ldots\ldots(12)$$

وتقرن القيمة المحسوبة مع قيمة $F_{(r1,r2,\alpha)}$ بشان اتخاذ القرار بقبول أو رفض فرضية العدم التي تنص على انه لا توجد فروقات معنوية بين التباينات لداخل المجموعات وبين المجموعات .

اختبار F

وبالنظرلصعوبة الحصول على القيمة الجدولية لعدم توفرالجداول الخاصة بمعيار ولكس، وضع معيار بديلاً آخرمنقبل Rao ، وهذاالمعيارهومقياس F وصيغته الرياضية هي [5]:

$$F = \frac{1-\Lambda^{1/5}}{\Lambda^{1/5}} * \frac{ms-2\lambda}{p(k-1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(13)$$

بدرجات حرية $df_1 = p(k-1)$

$$df_1=ms - 2\lambda$$

$$m = N - {}^{1}\!/_{2}\,[\,P + k]$$

$$S = \left[\frac{p^2(1-k)^2-4}{(1-k)^2+p^2-5}\right]^{1/2}$$

$$\lambda = \frac{p(k-1)-2}{4}$$

التمييزالقويم

إن الفكرة الأساسية في التحليل القويم هي تشكيل تركيبتين خطيتين واحدة لمجموعة(X) من المتغيرات والثانية لمجموعة(Y)من المتغيرات دون الحاجة إلى تحديد مجموعة المتغيرات المستقلة ومجموعة المتغيرات المعتمدة ،وذلك لقدرة التحليل القويم على قياس العلاقة بين المجموعتين أو إعطاء كل واحدة منها القدرة على التنبؤ بالمجموعة الأخرى ومن ثم اختز الهذه العلاقات إلى اقل عدد ممكن من المتغيرات حيث إن كل زوج من المتغيرات القويمة يمثلها معامل ارتباط بسيط ويمكن إعطاء فكرة عن تحليل الارتباط القويم فإذا كان لدينا (n) من المشاهدات و(p) تمثل المجموعة الأولى من المتغيرات ، (q) تمثل المجموعة الثانية من المتغيرات بحيث إن $n \geq p+qP \geq q$ [1] :

أي أن

$$X= x_{ij} \quad i=1,2,\ldots\ldots n \quad j= 1,2,\ldots..,p$$

$$Y= y_{ik} \quad i=1,2,\ldots\ldots n \quad k= 1,2,\ldots.....,p$$

إن التراكيب الخطية لهاتين المجموعتين هي $u=c'X$ و $v = d'Y$

وان c' وd' متجه الأوزان الخاصة بكل تركيبة خطية لكلتا المجموعتين على التوالي، حيث إن عدد الأزواج من التراكيب الخطية الناتجة من التحليل القويم يكون مساويًا إلى عددالمتغيرات في المجموعة الأقل $R=\min(p,q)$ وان R يمثل عدد الأزواج من التراكيب الخطية .ان المجموعتين من Yالمتغيرات التي ذكرت سابقا يمكن أن توصف من خلال التوزيع المشترك (joint distr).

لنأخذ :

$E(\underline{X}_1)=\underline{\mu}_1$, $E(\underline{X}_2)=\underline{\mu}_2$

$cov(\underline{X}_1)=\Sigma_{11}$,$cov(\underline{X}_2)=\Sigma_{22}$ ……..……(14)

$cov(\underline{X}_1,\underline{X}_2)=\Sigma_{12}=\Sigma_{21}$

علية يمكن تعريف المتجه $\underline{X}$ كالاتي :

$$\underline{X}_{((p+q)*1)}=\begin{bmatrix}\underline{X}_1\\ \cdots\\ \underline{X}_2\end{bmatrix}=\begin{bmatrix}x_{11}\\ x_{12}\\ \vdots\\ x_{1p}\\ \cdots\\ x_{21}\\ x_{22}\\ \vdots\\ x_{2p}\end{bmatrix} \qquad \text{……..(15)}$$

وله متجه المتوسطات :

$$\underline{\mu}_{((p+q)*1)}=E(\underline{X})=\begin{bmatrix}\underline{X}_1\\ \cdots\\ \underline{X}_2\end{bmatrix}=\begin{bmatrix}\underline{\mu}_1\\ \cdots\\ \underline{\mu}_2\end{bmatrix} \text{……(16)}$$

ومصفوفة التباينات والتغايرات المشتركة :

$$\sum_{(p+q)*(p+q)} = E\left(\underline{X}-\underline{\mu}\right)(\underline{X}-\underline{\mu})'$$

$$=\begin{bmatrix}E(\underline{X}-\underline{\mu})(\underline{X}-\underline{\mu})' & \vdots & E(\underline{X}-\underline{\mu})(\underline{X}-\underline{\mu})'\\ \cdots & \cdots & \cdots\\ E(\underline{X}-\underline{\mu})(\underline{X}-\underline{\mu})' & \vdots & E(\underline{X}-\underline{\mu})(\underline{X}-\underline{\mu})'\end{bmatrix}$$

$$=\begin{bmatrix}\Sigma_{11} & \vdots & \Sigma_{12}\\ \cdots & \cdots & \cdots\\ \Sigma_{21} & \vdots & \Sigma_{22}\end{bmatrix} \text{……(17)}$$

وإذا أخذنا التراكيب الخطية

$U=\underline{\acute{a}}\underline{X}_1$, $V=\underline{\acute{b}}\underline{X}_2$ ……..……(18)

لزوج متجهات المعاملات $\underline{a}$, $\underline{b}$ علية يمكن ايجاد :

$Var(U)=\underline{\acute{a}}cov(\underline{X}_1)\underline{a}=\underline{\acute{a}}\Sigma_{11}\underline{a}$

$Var(V)=\underline{\acute{b}}cov(\underline{X}_2)\underline{b}=\underline{\acute{b}}\Sigma_{11}\underline{b}$ ……..……(19)

$Var(U,V)=\underline{\acute{a}}cov(\underline{X}_1,\underline{X}_2)\underline{b}=\underline{\acute{a}}\Sigma_{12}\underline{b}$

ومنها يمكننا أن نوجد الارتباطات بين U,V أي أن

$\max_{a,b} corr(U_1,V_1)=e_1^*$

وهي تمثل أول زوج من المتغيرات القويمة وهو عبارة عن زوج التراكيب الخطية أي أن

(U_1,V_1) الذي يعظم الارتباط في المعادلة (19):

$U_1=\underline{e}_2'\,\Sigma_{11}^{-1/2}\underline{X}_1$

$$V_1=f_2'\sum_{22}^{-1/2}\underline{X}_2$$

تعظم

$\max_{e,f} corr(U_2,V_2)=e_2^*$

وهذا الزوج لا يرتبط بالزوج الأول .

وهكذا نجد إن الزوج k من المتغيرات القويمة حيث k=3,…..,p هو :

$U_k=\underline{e}_k^{'}\,\Sigma_{11}^{-1/2}\,\underline{X}_1$

$$V_k = f_k^{'}\sum_{22}^{-1/2}\underline{X}_2$$

يعظم نسبة الارتباط في المعادلة(19) أي أن :

$\max_{e,f}\mathrm{corr}(U_2, V_2) = e_2^*$

مع العلم بان هذه التراكيب الخطية لا ترتبط مع السابقة لها وهي 1,2,….,k-1 من المتغيرات القويمة .

ونخلص إلى إن $e_1^{*2} \geq e_2^{*2} \geq \cdots \geq e_p^{*2}$ هي عبارة عن القيم المميزة للمصفوفة $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{11}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ والمتجهات المميزة المرتبطة بها $e_1, e_2, \ldots, e_p$ وكذلك $e_1^{*2}, e_2^{*2}, \ldots, e_p^{*2}$ هي أيضا القيم المميزة للمصفوفة $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{11}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$ بالمتجهات المميزة $f_1, f_2, \ldots, f_p$.

ارتباطات العينة القويمة ومتغيرات العينة

العينة العشوائية التي تتكون من n من المشاهدات والمتغيرات (p+q) لـــــــــ x_2 و x_1 يمكن ان تجمع في مصفوفة البيانات التالية ((n(p+q)) [1]:

$$x_1=\lfloor x_{(1)} \;\vdots\; x_{(2)} \rfloor$$

$$=\begin{bmatrix} x_{11(1)} & x_{12(1)} & \cdots & x_{1p(1)} & \vdots & x_{11(2)} & x_{12(2)} & \cdots & x_{1q(1)} \\ x_{21(1)} & x_{22(1)} & \cdots & x_{2p(1)} & \vdots & x_{21(2)} & x_{22(1)} & \cdots & x_{2q(2)} \\ & & & & \vdots & & & & \\ x_{n1(1)} & x_{11(1)} & \cdots & x_{np(1)} & \vdots & x_{n1(2)} & x_{n2(2)} & \cdots & x_{nq(1)} \end{bmatrix}$$

$$=\begin{bmatrix} x_{1(1)}{}' & \vdots & x_{1(2)}{}' \\ \vdots & \vdots & \vdots \\ x_{n(1)}{}' & \vdots & x_{n(2)}{}' \end{bmatrix} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(20)$$

ومتجه متوسطات العينة يكون :

$$\underline{\overline{x}}_{[(p+q)*1]}=\begin{bmatrix} \underline{\overline{x}}_1 \\ \cdots \\ \underline{\overline{x}}_2 \end{bmatrix}$$

حيث :

$$\underline{\overline{x}}_1=\frac{1}{n}\sum_{j=0}^{n}\underline{x}_{j(1)} \quad , \quad \underline{\overline{x}}_2=\frac{1}{n}\sum_{j=0}^{n}\underline{x}_{j(2)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(21)$$

وبالمثل مصفوفة التباين والتغاير المشترك للعينة يكون :

$$S_{(p+q)*(p+q)}=\begin{bmatrix} S_{11} & \vdots & S_{12} \\ \cdots & \cdots & \cdots \\ S_{21} & \vdots & S_{22} \end{bmatrix}$$

حيث :

$$S_{KL} = \frac{1}{n-1}\sum_{j=0}^{n}(x_{j(k)} - \underline{\overline{x}}_{(k)})\left(\left(x_{j(L)} - \underline{\overline{x}}_{(L)}\right)\right)^{'};\mathrm{K,L=1,2} \quad \ldots\ldots\ldots(22)$$

وتكون العينة التراكيب الخطية :

$$\hat{u}= \underline{\hat{a}}X_1 \quad , \quad \hat{v}= \underline{\hat{b}}X_2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(23)$$

الارتباطات التالية :

$$r_{\hat{u},\hat{v}}=\frac{\underline{\hat{a}}'S_{12}\underline{\hat{b}}}{\sqrt{\underline{\hat{a}}'S_{12}\underline{\hat{a}}}\sqrt{\underline{\hat{b}}'S_{12}\underline{\hat{b}}}} \quad \ldots\ldots\ldots\ldots\ldots(24)$$

وأول زوج من متغيرات العينة القويمة هو زوج التراكيب الخطية $\hat{u},\hat{v}$الذي يعظم النسبة (2-99)بصورة عامة الزوج رقم k من متغيرات العينة القويمة هو زوج التراكيب الخطية $\hat{u}_k,\hat{v}_k$ الذي سنعظم النسبة (24). وهو غير مرتبط مع الأزواج k=1,2,…k-1 وارتباط العينة بين $\hat{u}_k,\hat{v}_k$ يعرف بـــــ k^{th} ارتباط العينة القويم .كما نجد أن ارتباطات العينة القويمة $r_1 \geq r_2 \geq \cdots \geq r_p$ هي الجذور التربيعية الموجبه غير الصفرية للقيم التميزية للمصفوفة $S_{21}S_{11}^{-1}S_{12}S_{22}^{-1}$ والمتجهات $\underline{\hat{a}}_k,\underline{\hat{b}}_k$ تقابل k^{th} من زوج المتغيرات القويمه والتي تعطي من الحلول :

$$S_{22}^{-1}S_{21}S_{11}^{-1}S_{12}\underline{\hat{b}}_k = r_k^2\underline{\hat{b}}_k$$
$$S_{11}^{-1}S_{12}S_{22}^{-1}S_{21}\underline{\hat{a}}_k = r_k^2\underline{\hat{a}}_k \ldots\ldots\ldots\ldots\ldots\ldots\ldots(25)$$

حيث r_k^2 تحقق

$$\left|S_{12}S_{11}^{-1}S_{21}\underline{\hat{a}}_k - r_k^2S_{22}^{-1}\right|=0$$

ويبقى السؤال المهم هنا هو(ما هو عدد المتغيرات القويمة التي يمكن أن تصف وتمييز البيانات بشكل كاف ؟) وإجابة هذا السؤال تكمن في معرفة عدد r من الارتباطات القويمة التي تختلف معنويا عن الصفر ، وبافتراض ان العينات تتوزع وفق التوزيع الطبيعي متعدد المتغيرات والاختبار المتتابع لفرضية العدم :

$$H_o^{(j)}: r_1 = r_{j+1} = \ldots\ldots\ldots r_p \quad ; j=1,2,\ldots\ldots,r$$

وكلما كانت النتيجة هي رفض فرضية العدم أعلاه فإننا نستمر في اختبار الفرضية التي تليها، ونتوقف عندما تقبل فرضية العدم $H_o^{(j)}$. ولنفترض ان فرضيات العدم $H_o^{(1)}, H_o^{(2)} \ldots, H_o^{(j-1)}$ كلها رفضت ولكن $(H_o^{(j)}; j = 1,2, \ldots r+1)$ كانت مقبولة . وهذا الافتراض يعني انه فقط j-1 من المتغيرات القويمة هي كافية لوصف وتمييز البيانات .وهنالك بعض الخصائص بالنسبة للقيم المميزة هي

$$var(\hat{u}_k)=var(\hat{v}_k)=i$$
$$cov(\hat{u}_k,\hat{u}_e)=corr(\hat{u}_k,\hat{u}_e)=0 \ , k\neq e$$
$$cov(\hat{v}_k,\hat{v}_e)=corr(\hat{v}_k,\hat{v}_e)=0 \ , k\neq e$$
$$cov(\hat{u}_k,\hat{v}_e)=corr(\hat{u}_k,\hat{v}_e)=0, k\neq e, \forall k,e=1,\ldots,p \ \ldots\ldots\ldots\ldots\ldots(26)$$

حساب الأوزان القويمة canonical weights

أنأهمية تأثيرالمجموعة الأولى وتأثرالمجموعة الثانية باستخدام الأداة الإحصائية المسماة بالارتباط القويم تعتمد على الأوزان القويمة والتي توضح قوة تأثير أي من متغيرات المجموعة الأولى نسبة إلى متغيرات نفس المجموعة الأخرى وأيضا قوة تأثرأي من متغيرات المجموعة الثانية نسبة إلى متغيرات نفس المجموعة الأخرى[1].

ولحساب الأوزان القويمة تستخدم معادلة المتجه المميزالآتية:

$$(M-\lambda I)d = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots(27)$$

حيث أن

$$M= S_{yy}^{-1}S_{yx}\,S_{xx}^{-1}S_{xy}\ldots\ldots\ldots\ldots\ldots(28)$$

بحيث لوكان $P \leq q$ فان المعادلةالتي يستخرج منها الارتباط القويم هي:

$$S_{xx}^{-1}S_{xy}\,S_{yy}^{-1}S_{yx} - \lambda I \ \ldots\ldots\ldots\ldots\ldots\ldots(29)$$

بحيث لوكان $q \leq p$ فان المعادلة التي يستخرج منها الارتباط القويم هي:

$$S_{yy}^{-1}S_{yx}\,S_{xx}^{-1}S_{xy} - \lambda I \ \ldots\ldots\ldots\ldots\ldots\ldots(30)$$

حيث أن :-

$S_{xx}\equiv$ مصفوفة ال Var-Cov للمستويات في المجموعة X

$S_{yy}\equiv$ مصفوفةال Var-Cov للمستويات في المجموعة Y

$S_{xy}\equiv$مصفوفةال Var-Cov للمستويات في المجموعة XY

كذلك فانه بالإمكان الحصول على الارتباط القويم باستخدام مصفوفة الارتباطات كما يمكن الحصول عليه من مصفوفة التباين – التباين المشترك.

اختبارمعنوية الارتباط القويم

ليست كل الارتباطات التي تم الحصول عليها هي ارتباطات مفيدة وذات أهمية بالمعنى الإحصائي فهناك ارتباطات لاتأثيرلها أساسا في التطبيق المستخدم، لذا فاختبارمعنوية الارتباط القويم هوالجزءالمكمل للارتباط القويم حيث أوضح هذه الفكرة العالم (Bartlett) في عام 1938م حيث وض عدالة خاصة لاختبار معنوية الجذور القويمة وكما مبينة في أدناه:

إن الفرضية الخاصة بالاختبارهي[6]

$H_0 : R_{XY} = 0$
$H_1 : R_{XY} \neq 0$

ويتم استخدام احصاءة χ^2 التي تحسب على وفق المعادلةالآتية:

$$\chi^2_{cal} = [-n + 0.05(p + q + 3)]\log w \quad \text{.....(31)}$$

حيث إن w تسمى احصاءة Wilk

$$\text{Wilks} \leftarrow w = \prod_{Zzi}^{r}(1 - R_{cz}^2) \quad \text{.....(32)}$$

R^2_{cz}:مربع معامل الارتباط القويم

فإذا كان الارتباط القويم الأول غيرمعنوي، فأن بقية الارتباطات القويمة الأخرى تكون غيرمعنوية أيضا . ومن هنا نستنتج أن اختبار أهمية الجذور المعنوية هو شئ ضروري جدًا ولابد منه لأنه يخفف من العبء التعبيري على الباحث وذلك من خلال استبعاد الجذورالغيرمعنوية والتركيز على الجذورالمعنوية في إتمام متطلبات هذا التحليل..

مجموع المربعات التنبؤي Prediction Sum of Square (PRESS)

نحسب مجموع المربعات التنبوي بوضع الملاحظة الأولى ثم نقوم بوضع الملاحظة الثانية جانباً مع تقييم المعامل مره اخرى . نقوم بابعاد الملاحظات كل على حدا ، وهذا يمثل ان المرشح مناسباً (n) ومرات الاستجابة الملغاة تقيم كل مرة ، معتمدة على نتيجة الأخطاء التنبؤية في (n) بالضغط المتبقي

$$y_i - \hat{y}_{i.-1} = e_{i,-1} \quad (i=1,2,\ldots,n) \quad \text{.....(33)}$$

هذا الضغط المتبقي هو أخطاء تنبؤية حقيقية تكون فيه $\hat{y}_{i.-1}$ مستقلة عن $y_{i.}$. ولذلك بهذه الطريقة ، الملاحظة ان y_i لا يمكن استخدامها متزامنة للمطابقة ونموذج للتقييم ، هذا يعتبر اختباراً حقيقياً للصلاحية. التنبؤ $\hat{y}_{i.-1}$ يعتبر وظيفة ارتدادية تقدر في $x=x_{i,}$ لكن y_i توضع جانباً ولا تستخدم في الابقاء على المعاملات . قومياً او وطنياً

$$\hat{y}_{i.-1} = \acute{x}_i b_{-i} \quad \text{.....(34)}$$

حيث تمثل b_{-i} مجموعة من المعاملات أحصيت بدون استخدام الملاحظة في ith ولذا سوف يمتلك كل نموذج مرشح (n) حفظ متبقيات متعلقاً بها . وحفظ (قيمة تنبؤية للمربعات) تعرف كالأتي :-

$$\text{PRESS} = \sum_{i=1}^{n}(y_i - \hat{y}_{i.-1})^2 = \sum_{i=1}^{n}(e_{i,-1})^2 \quad \text{.....(35)}$$

الجانب التطبيقي

النتائج

تحديد المتغيرات ذات التأثير المعنوي باستخدام التمييز القويم

الجدول 1 : نتائج اختبارwilk'sLambda واختبار F للمتغيرات التمييزية:

		F	Wilk's Lambda
.1	السنة X_{01}	2.756	0.966
.2	النوع X_{02}	2.478	0.970
.3	العمرX_{03}	2.434	0.970
.4	القبيلة X_{04}	1.308	0.984
.5	المهنة X_{05}	3.046	0.963
.6	السكنX_{06}	0.338	0.996
.7	الحالةالاجتماعيةX_{07}	0.181	0.998
.8	عددالأطفال X_{08}	0.413	0.995
.9	مستوىالدخل X_{09}	2.653	0.968

من الملاحظ في الجدول (1)إن المتغيرx_{05} والذي يمثل المهنة (X_{05}) هو المتغير الداخل الأول حيث كانت له اقل قيمهwilk's Lambda(0.963) واكبر قيمه F (3.046) . ثم نستخرج F للمتغير الخارج والتي تساوي F للمتغير الداخل في هذه الخطوةأي تساوي (3.046) . ثم نجري الخطوات الأخرى باستخراج wilk's Lambda ، F لكل متغير غير داخل بنفس الطريقةالسابقة (مع تغيير في المصفوفات المستخدمة في كل خطوه) وكذلك Fللمتغير الخارج بالنسبة للمتغيرات الداخلة كما ونلاحظ أن المتغير الداخل الأخير هوX_{08}ويمثل القبيلة والجدول رقم (2)يبين خلاصة نتائج الخطوات المشار إليهاأعلاه .

الجدول 2: إدخال المتغيرات التمييزية حسب أهميتها:

Step	Variable enter	Wilks' Lambda and its degrees of freedom	F to remove and its degrees of freedom
1.	القبيلة X_{04}	0.984 (6,4,317)	1.308 (24, 0.267)
2.	النوع X_{02}	0.970 (4,4,317)	2.478 (16, 0.044)
3.	العمر X_{03}	0.970 (5,4,317)	2.434 (20, 0.047)
4.	مستوىالدخل X_9	0.968 (3,4,317)	2.653 (12, 0.033)
5.	السنة X_{01}	0.966 (2,4,317)	2.756 (8, 0.028)
6.	المهنة X_{05}	0.963 (1,4,317)	3.046 (4, 0.017)

تكوين دوال التمييز ثم التصنيف

1 - تكوين دوال التمييز

بعد تحديد المتغيرات ذات التأثير المعنوي واستخراج المصفوفاتتستخرج مصفوفة المعاملات B حسب المعادلة (1) ثم نستخرج المصفوفة R حيث:

$$R = S^{-1} B$$

بعد تحديد القيم العينية والتي توضح الأهمية النسبية لكل داله من دوال التمييز ، نستخرج النسبة المئوية للتباين الذي توضحه كل دالة من الدوال وذلك حسب المعادلة وكانت النسب كما في الجدول رقم (3).

الجدول 3: النسب المئوية للتباين لدوال التمييز

Function	Percent of Variance
1	67.9
2	28.0
3	3.0
4	1.1

ونلاحظ من الجدول (3) أن الدالتين الأولىوالثانية تشرحان 95.9 من التباين بعد استخراج القيم العينية نعوضها نستخرج معاملات التمييز α،s كما موضحه في الجدول أدناه .

الجدول 4: معاملات دوال التمييز القويم

Function Coff	1	2	3	4
α_1السنه	0.093	0.059	0.034	-2.265
α_2النوع	-0.276	0.826	-0.406	-0.112
α_3عمر	0.582	0.571	0.430	0.167
α_4القبيله	0.439	-0.291	-0.143	-0.099
α_5المهنه	-0.499	-0.214	-0.567	1.594
α_9مستوىالدخل	-0.336	-0.033	1.052	0.719

بعد تحديد معاملات التمييز يمكن كتابة دوال التمييز كلاتي :

الدالة الأولى : وتشرح نسبة 67.9% من التباين :

$$Z'_1 = 0.093\alpha_1 - 0.276\alpha_2 + 0.582\alpha_3 + 0.439\alpha_4 - 0.499\alpha_5 - 0.336\alpha_9$$

الدالةالثانية : وتشرح نسبة 28.0% من التباين :

$$Z'_2 = 0.059\alpha_1 + 0.826\alpha_2 + 0.571\alpha_3 - 0.291\alpha_4 - 0.214 - 0.033\alpha_9$$

الدالة الثالثة : وتشرح نسبة 3.0% من التباين :

$$Z'_3 = 0.034\alpha_1 - 0.406\alpha_2 + 0.430\alpha_3 - 0.143\alpha_4 - -0.567\alpha_5 + 1.052\alpha_9$$

الدالة الرابعة : وتشرح نسبة 1.1% من التباين :

$$Z'_4 = -2.265\alpha_1 - 0.112\alpha_2 + 0.167\alpha_3 - 0.099\alpha_4 + 1.594\alpha_5 + 0.719\alpha_9$$

التصنيف Classiffication

نستخرج المعاملاتα_i's حسب المعادلة (6) وكذلك الثابت C_iكما في المعادلة (5) وتكون النتائج كما يلي :

الجدول 5: معاملات دوال التصنيف ClassificationDiscriminant Function Coefficients

Function Coff	1	2	3	4
α_{i1}السنه	0.041	0.026	0.015	-0.994
α_{i2}النوع	-0.562	1.680	-0.825	-0.227
α_{i3}عمر	0.774	0.759	0.572	0.222
α_{i4}القبيله	0.240	-0.159	-0.078	-0.054
α_{i5}المهنه	-0.131	-0.056	-0.149	0.418
α_{i9}مستوىالدخل	-0.450	-0.045	1.409	0.963
(Constant)	0.193	-3.063	-1.178	-0.060

وتكون المعادلة كالأتي:-

$$Z'_1=0.041\alpha_{i1}-0.562\alpha_{i2}+0.774\alpha_{i3}+0.240\alpha_{i4}-0.131\alpha_{i5}-0.450\alpha_{i9}+0.193+\ln p_i$$
$$Z'_2=0.026\alpha_{i1}+1.680\alpha_{i2}+0.759\alpha_{i3}\alpha_{i4}-0.159\alpha_{i5}-0.045\alpha_{i9}-3.063+\ln p_i$$
$$Z'_3=0.015\alpha_{i1}-0.825\alpha_{i2}+0.572\alpha_{i3}-0.078\alpha_{i4}-0.149\alpha_{i5}+1.409\alpha_{i9}-1.178+\ln p_i$$
$$Z'_4=-0.994\alpha_{i1}-0.227\alpha_{i2}+0.222\alpha_{i3}-0.054\alpha_{i4}+0.418\alpha_{i5}+0.963\alpha_{i9}-0.060+\ln p_i$$

وبعد استخراج الدوال أعلاه نستخرج الاحتمالات اللاحقة Posterior Probability كما في المعادلة (7)حيث يكون لكل قيمه يراد تصنيفها أربعة قيم للاحتمالات اللاحقة وكل قيمه توضح احتمال وقوع المفردة في المجموعة المقابلة ، حيث تصنف إلى المجموعة التي تقابل اكبر قيمة احتمال لاحق .

بناء دوال التمييز باستخدام التمييز القويم Canonical Discriminate Function

بناء دوال التمييز بواسطة التمييز القويم يتطلب أولاإيجاد المتوسطات والانحرافات المعيارية للمجموعات وكذلك نتائج اختبار Wilk's lambda واختبار F للمتغيرات التمييزية ولقد تم التوصل إليها سابقاًفي الجدول (1) . فمن الجدول (1)نلاحظ أن متغير المهنة (X_{06})كانت اكبر قيمة لـــــــ F (3.046) واقل قيمة Wilk's lambda (0.963) عليه فهو المتغير الأكثرمساهمه في التمييز بين المجموعات [2] .

تكوين دوال التمييز ثم التصنيف

تكوين دوال التمييز

بعد تحديد المتغيرات ذات التأثير المعنوي واستخراج المصفوفات كما نستخرج مصفوفة المعاملات B حسب المعادلة (1) ثم نستخرج المصفوفة R حيث يوضح ذلك الجدول رقم (6).

الجدول 6 : نتيجة اختبار Box's M :

Box's M		158.535
F	Approx.	0.812
	df1	180
	df2	64373.551
	Sig.	0.970

بعد تحديد القيم العينية والتي توضح الأهمية النسبية لكل داله من دوال التمييز نستخرج النسبة المئوية للتباين الذي توضحه كل دالة من الدوال حسب المعادلة(2) وكذلك النسب كما في الجدول رقم (7).

الجدول 7: النسب المئوية للتباين لدوال التمييز القويم

Percent of Variance for Discriminant Function

Function	Percent of Variance
1	62.8
2	31.3
3	3.6
4	2.4

ونلاحظ من الجدول (7) أن الدالتين الأولى والثانية تشرحان 94.1 من التباين .وهي اقل من النسبة التي تشرحها الدالتين في الجدول (4) وهذا يدل على ان تمييز المراحل يعطي نسبه اكبر للدوال في التمييز من الذي تعطيه طريقة التمييز القويم .وبعد استخراج القيم العينية نستخرج قيم معاملات التمييز كما موضحة بالجدول رقم (8).

الجدول (8):معاملات دوال التمييز القويم *Canonical Discriminate Function Coefficients*

Function Coff	1	2	3	4
α_1السنة	0.083	-0.038	0.661	-1.416
α_2النوع	-0.294	0.731	-0.180	-0.382
α_3 العمر	0.533	0.585	0.421	0.104
α_4 القبيلة	0.451	-0.247	-0.094	-0.059
α_5 المهنة	-0.513	-0.107	-0.810	0.818
α_6 السكن	0.169	0.186	-0.295	0.644
α_7الحالةالاجتماعية	-0.010	-0.069	0.504	-0.030
α_8 عددالأطفال	0.135	-0.423	0.030	-0.132
α_9 مستوىالدخل	-0.342	-0.041	0.635	0.850

بعد تحديد معاملات التمييز يمكن كتابة دوال التمييز كالأتي :

الدالة الأولى : وتشرح نسبة 62.8% من التباين :

$$Z'_1=0.083\alpha_1-0.294\alpha_2+0.533\alpha_3+0.451\alpha_4-0.513\alpha_5+0.169\alpha_6-0.010\alpha_7+0.135\alpha_8-0.342\alpha_9$$

الدالة الثانية : وتشرح نسبة 31.3% من التباين :

$$Z'_2=-0.038\alpha_1+0.731\alpha_2+0.585\alpha_3-0.247\alpha_4-0.107\alpha_5+0.186\alpha_6-0.069\alpha_7-0.423\alpha_8-0.041\alpha_9$$

الدالة الثالثة : وتشرح نسبة 3.6% من التباين :

$$Z'_3=0.661\alpha_1-0.180\alpha_2+0.421\alpha_3-0.094\alpha_4-0.810\alpha_5-0.295\alpha_6+0.504\alpha_7+0.030\alpha_8-0.635\alpha_9$$

الدالة الرابعة : وتشرح نسبة 2.4% من التباين :

$$Z'_4=-1.416\alpha_1-0.382\alpha_2+0.104\alpha_3-0.059\alpha_4+0.818\alpha_5+0.644\alpha_6-0.030\alpha_7-0.132\alpha_8+0.850\alpha_9$$

الجدول 9: مصفوفة الارتباط بين متغيرات الدراسة *Structure Matrix*

Function Coff	1	2	3	4
x_{05}المهنة	-0.635	-0.055	0.150	0.327
x_{01} السنة	-0.600	-0.035	0.417	0.103
x_{09}مستوىالدخل	-0.581	-0.008	0.535	0.282
x_{04} القبيلة	0.390	-0.204	-0.161	-0.217
x_{02}النوع	-0.336	0.652	-0.207	-0.402
x_{03}العمر	0.442	0.486	0.468	0.241
x_{08}عددالأطفال	0.075	-0.314	0.124	-0.061
x_{07}الحالةالاجتماعية	0.101	-0.093	0.416	0.054
x_{06}السكن	0.039	0.225	-0.149	0.683

من الجدول (9) أعلاه يتضح لنا ترتيب العوامل على حسب أهميتها التمييزية . فنلاحظ أن المتغير x_{05} ويمثل المهنة هواهم المتغيرات في التمييز ، ثم بالترتيب الموضح حتى آخر متغير x_{06} ويمثل العمر.

الجدول 10:معاملات دوال التمييز القويم *Canonical Discriminate Function Coefficients*

Function Coff	1	2	3	4
α_{i1} السنة	0.037	-0.017	0.290	-0.621
α_{i2} النوع	-0.597	1.485	-0.367	-0.776
α_3 العمر	0.708	0.778	0.559	0.139
α_4 القبيلة	0.247	-0.135	-0.052	-0.033
α_5 المهنة	-0.134	-0.028	-0.213	0.215
α_6 السكن	0.066	0.073	-0.115	0.252
α_7 الحالةالاجتماعية	-0.011	-0.075	0.548	-0.032
α_8 عددالأطفال	0.243	-0.758	0.054	-0.236
α_9 مستوىالدخل	-0.458	-0.055	0.850	1.138
(Constant)	-0.162	-2.032	-2.336	-0.406

وتكون المعادلات كما في المعادلة

$$Z'_i=\alpha_{i1}x_1+\alpha_{i2}x_2+\alpha_{i3}x_3+\alpha_{i4}x_4+\alpha_{i5}x_1+\alpha_{i6}x_{6+}\alpha_{i7}x_7+\alpha_{i8}x_8+\alpha_{i9}x_9+\ln p_i$$

$$Z'_1=0.037\alpha_i-0.597\alpha_{i2}+0.708\alpha_{i3}+0.247\alpha_{i4}-0.134\alpha_{i5}+0.066\alpha_{i6}-0.011\alpha_{i7}+0.243\alpha_{i8}-0.458\alpha_{i9}-0.162$$

$$Z'_2=-0.017\alpha_{i1}+1.485\alpha_{i2}+0.778\alpha_{i3}-0.135\alpha_{i4}0.028\alpha_{i5}+0.073\alpha_{i6}-0.075\alpha_{i7}-0.758\alpha_{i8}-0.055\alpha_{i9}-2.032$$

$$Z'_3=0.290\alpha_{i1}-0.367\alpha_{i2}+0.559\alpha_{i3}-0.052\alpha_{i4}-0.213\alpha_{i5}+-0.115\alpha_{i6}+0.548\alpha_{i7}+0.054\alpha_{i8}-0.850\alpha_{i9}-2.336$$

$$Z'_4=-0.621\alpha_{i1}-0.776\alpha_{i2}+0.139\alpha_{i3}0.052\alpha_{i4}+0.215\alpha_{i5}+0.252\alpha_{i6}-0.032\alpha_{i7}-0.236\alpha_{i8}+1.138\alpha_{i9}-0.406$$

وبعد استخراج الدوال أعلاه نستخرج الاحتمالات اللاحقة Posterior Probability كما في المعادلة (7)حيث يكون لكل قيمه يراد تصنيفها أربعة قيم للاحتمالات اللاحقة وكل قيمه توضح احتمال وقوع المفردة في المجموعة المقابلة ، حيث تصنف إلى المجموعة التي تقابل اكبر قيمة احتمال لاحق .

تحديد المتغيرات ذات التأثير المعنوي باستخدام مجموع المربعات التنبؤي

نلاحظ من الجدول (11) أن المتغيرات ذات دلالة إحصائية أي العلاقة بين المتغيرات كافية بقيمة لقيصر 0.559 و هذه القيمة اكبر 0.05 هذا يعني إن المتغيرات لها تأثير معنوي على المرضى المصابين بمرض الفشل الكلوي المزمن .

الجدول 11 : اختبار كفاية العينة(K.M.O)

قيمة قيصر	قيمة مربع كاي	درجة الحرية	القيمة الاحتمالية
0.559	353.271	28	0.000

الجدول 12 : اختبار كلومجروف - سيمرنوف

	قيمة كلومجروف	درجة الحرية	القيمة الاحتمالية
النوع x_2	0.375	322	0.000
عمر x_3	0.293	322	0.000
القبيلة x_4	0.235	322	0.000
المهنة x_5	0.159	322	0.000
السكن x_6	0.221	322	0.000
الحالةالاجتماعية x_7	0.264	322	0.000
عددالأطفال x_8	0.447	322	0.000
مستوىالدخل x_9	0.280	322	0.000

يتضح من نتائج التحليل الإحصائي لاختبار كلومجروف – سيمرنوف أن القيم الاحتمالية اصغر من 5% بالتالي فإننا نرفض العدم القائل بان المتغيرات لا تتبع التوزيع الطبيعي ، ونقبل الفرض البديل القائل بان توزيع المتغيرات يتبع التوزيع الطبيعي .هذا يعني إن المتغيرات لها تأثير معنوي على المرضى المصابين بمرض الفشل الكلوي المزمن .

الجدول 13 : معامل الالتواء للمتغيرات

	النوع	عمر	القبيلة	المهنة	السكن	الحالةالاجتماعية	عددالأطفال	مستوىالدخل
الوسط الحسابي	1.43	1.90	2.74	6.99	4.11	1.88	1.31	1.75
الوسيط	1.00	2.00	2.00	7.00	4.00	2.00	1.00	2.00
المنوال	1	2	1	2	4	2	1	1
الالتواء	0.264	0.771	0.594	0.010	1.086	0.956	1.596	0.453

يتضح من الجدول (13) وبما إن الوسط = الوسيط = المنوال هذا يعني ان الالتواء متماثل اي توزيع المتغيرات اعتدالي . وهذا يدل على إنها تتبع للتوزيع الطبيعي .

معادلة الانحدار الخطي لمتغيرات الدراسة حسب أهميتها

$$\hat{y}=3.03-0.005X_{01}+0.327X_{02}+0.236X_{03}+0.0005X_{05}-0.0186X_{05}-0.117X_{09}$$

جدول رقم14 : يوضح معاملات نموذج الانحدار والثوابت واختبار T للمتغيرات

Predictor	Coef	SE Coef	T	P
Constant	3.0264	0.3726	8.12	0.000
النوع x_2	0.3272	0.1628	2.01	0.045
العمر x_3	0.2357	0.1064	2.22	0.027
القبيلة x_4	-0.00028	0.04395	-0.01	0.995
المهنة x_5	-0.01856	0.03296	-0.56	0.574
مستوى الدخل x_9	0.1171	0.1692	0.69	0.043

يتضح من الجدول (14) قيمة الاحتمال لكل من الجزء الثابت من لانحدار B_0 تساوي 0.000 ومعامل انحدار المتغير الأول x_1 تساوي 0.043 ومعامل انحدار المتغير الثاني x_2 تساوي 0.045 ومعامل انحدار المتغير الثالث x_1 تساوي 0.027 وهي اقل من مستوى معنوية 5% وبالتالي نرفض الفرض العدم القائل أن المقدر الثابت في نموذج الانحدار غير معنوي ، ونقبل الفرض البديل القائل أن المقدر الثابت في نموذج الانحدار معنوي.هذا يعني أن للمتغيرات تأثير معنوي في علاج الإصابة بمرض الفشل الكلوي .

أما بالنسبة ل x_4 ، x_5 نجد أن القيمة الاحتمالية في الحالتين اكبر من مستوى المعنوية 5% وبالتالي نقبل الفرض العدم ونرفض الفرض البديل بعدم معنوية هاذين المعاملين في معنوية الانحدار .

جدول 15: تحليل التباين واختبار F للمتغيرات:

Source	DF	SS	MS	F	P
Regression	5	21.429	4.286	2.09	0.066
Residual Error	316	647.266	2.048		
Total	321	668.696			

يتضح من جدول تحليل التباين رقم (15) ان القيمة الاحتمالية للمتغيرات تساوي 0.066 وهي اقل من مستوى المعنوية 5% وبالتالي نرفض الفرض العدم القائل بان معاملات الانحدار غير معنوية ، ونقبل الفرض البديل القائل بان معاملات الانحدار معنوية.هذا يعني أن للمتغيرات تأثير معنوي في علاج الإصابة بمرض الفشل الكلوي المزمن .

جدول 16 : المقاييس الخاصة بالبواقي للمتغيرات

Obs	النوعالفشل	Fit	SE Fit	Residual	St Resid
7	2.00	1.0000	4.0152	0.2034	-3.0152	-2.13R
16	2.00	1.0000	4.0160	0.1900	-3.0160	-2.13R
41	2.00	1.0000	3.9964	0.2126	-2.9964	-2.12R
128	2.00	1.0000	3.9235	0.1862	-2.9235	-2.06R
313	2.00	5.0000	4.2350	0.4350	0.7650	0.56 X
314	1.00	3.0000	3.9072	0.4213	-0.9072	-0.66 X
315	2.00	4.0000	3.5279	0.3994	0.4721	0.34 X
316	2.00	4.0000	3.7633	0.3764	0.2367	0.17 X
317	2.00	5.0000	3.5268	0.4003	1.4732	1.07 X
318	1.00	5.0000	3.2001	0.3882	1.7999	1.31 X
319	2.00	5.0000	3.5268	0.4003	1.4732	1.07 X
320	1.00	2.0000	3.1993	0.4149	-1.1993	-0.88 X
321	2.00	1.0000	3.5265	0.4124	-2.5265	-1.84 X
322	1.00	3.0000	3.2004	0.3889	-0.2004	-0.15 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large leverage.
S = 1.43119 R-Sq = 3.2% R-Sq(adj) = 1.7%
PRESS = 671.576 R-Sq(pred) = 0.00%

يتضح من القيم أعلاه أن معامل التحديد المعدل يساوي 1.7% وهذا معناه أن المتغير المستقل(مراحل الفشل الكلوي) يفسر 1.7% من التغيرات التي تحدث في المغيرات التابعة والبواقي كما موضحة في الجدول أعلاه ، يرجع إلى عوامل أخرى منها الخطأ العشوائي .

الاستنتاجات

من خلال نتائج الدراسة فان هنالك بعض الاستنتاجات الخاصة بالموضوع وكما يلي:

1) الدالة المميزة المقدرة لها مقدرة على التمييز .
2) لقد أظهرت النتائج التي توصلنا اليها بان هناك تأثير معنوي عالي قياساً بالمتغيرات الأخرى للمتغير الذي يمثل المهنة من بين المتغيرات الكلية .
3) إن توزيع البيانات يتبع التوزيع الطبيعي .

المراجع

1 –عائده هادئ صالح .(2008م). استخدام التحليل المميز لتشخيص بعض أمراض العيون .
2 –د . عبد الكريم السويداء (2010م) المرشد الشامل لمرضى الفشل الكلوي.
3 –عفراء هاشم .(2011م). ببحث بعنوان تطبيق التحليل التمييزي لتصنيف الإصابة بسرطان الثدي لدى الإناث حسب مراحله المتعددة .
4 –الدكتور فريد الجاعوني ، والدكتور عدنان غانم.(2007 م). التحليل الإحصائي متعدد المتغيرات والتحليل التمييزي فتوصيف وتوزيع الأسر داخل الهيكل الاقتصادي والاجتماعي في المجتمع .
5 –بسام إبراهيم يونس .(2002م). " محاضر في تحليل متعدد المتغيرات " ملخصات للسنة الرابعة قسم الإحصاء التطبيقي بكلية العلوم جامعة السودان للعلوم والتكنولوجيا .
6 –دوخي الحنيطي وقبلان المجالي، وسعود الطيب، وحسن العثمان ، وامجد حراز.(2004 م). تمييز الأسر الفقيرة من غير الفقيرة في المناطق النائية التابعة لإقليم جنوب الأردن.
7 –مارفن أ .جولسون و وريتشارد ب.روكسي (2001م) بحث بعنوان " ارتباطات التعرض للخصومات
8 –نوح ، الطاهر ادم (2008 م) *استخدام الداله التمييزيه في تصنيف فئات الاجر* رسالة ماجستير مقدمة إلى جامعة السودان للعلوم والتكنولوجيا .
9 –هاشم ، هاله عثمان ، (2006 م) استخدام *الداله التمييزيه في تمييز الاصابه بمشكلات في القدمين لدى مرضى السكري"* رسالة ماجستير مقدمة إلى جامعة السودان للعلوم والتكنولوجيا .