استخدام التحليل المميز لتحديد أهمية العوامل المؤثرة على المصابين بمرض اليرقان للأطفال حديثي الولادة

Discriminate Analysis is used to determine the importance of distinctive factors affecting patients jaundice neonatal

طاهر ريسان دخيل
Taher Resan dakil
جامعة القادسية
Qadissiya university
قسم الاحصاء
Department of statistics
أستاذ مساعد
tahir_reisan@yahoo.com

زينب يوسف داود
Zainab Yousif Dawood
الجامعة المستنصرية
AL mustansiriya university
قسم الاحصاء
Department of statistics
مدرس
zainabassfor@yahoo.com

الخلاصة

من المعلوم ان من الأمراض الاكثر شيوعاً بين الأطفال حديثي الولادة هو مرض اليرقان الذي يحدث بنسبة عالية في المولودين قبل موعد الولادة الصحيح وذلك بسبب عجز الكبد عن تكوين إنزيم يسمى ((Glycony Itranfecs)) الذي يحتاج لعملية تصريف البيليروبين من الإفراز وعندما تزداد نسبته بالدم اكثر من 250 – 350 مليمول سوف يسبب تلف الدماغ ((Ternictaus)) وبالتالي عملية تبديل الدم تكون ضرورية.

استخدم التحليل التمييزي لدراسة البيانات المتوفرة من عدد الأطفال المصابين ونسبة أبو صفار بالدم ((اليرقان)) في الأطفال حديثي الولادة ومن أسباب هذا المرض تعرض ألام للأمراض خلال الحمل وتسمى هذه الحالة (Torch) أو اختلاف فصيلة الدم بين ألام والأب ، ولمعالجة هذا المرض يتم تبديل دم الطفل الوليد إذا كان نوعه مرضي و خلال ٢٤ ساعة ،أما إذا حدث خلال ٤٨ ساعة فحسب نسبة الاصابة أما زيادة الرضاعة إذا كانت نسبة الاصابة قليلة أو الأشعة الفوتية إذا كانت نسبة الاصابة عالية .

Abstract

It is well-known that the most wide spread diseases among infants is the janndice which occurs widly with predelivered fetus because of the disability of the liver to produce enzymes which called “Glycony Itranfecs” which needs to get rid of billion from excretion so when the rate of the billion in blood increases more than 250-350 milimole that will destroy the brain (Ternicataus) so the process of recharging the blood is necessary.

The discriminative analyses have been used to study the available data through the number of patients and its rate in blood. One of the reasons of this disease is that the mother being exposed during her pregnancy to diseases and this case is called “Torch” or the difference of the parents in blood group. To treat it ,on should change the blood babies if is diseased through 24 hours ,while if it happened through 48 hours ,so it depends on the rate either increasing the suckling if its rate was low or photo ray if its rate was high.

المقدمة

تصاب نسبة كبيرة من الاطفال حديثي الولادة باصفرار في لون الجلد والعينين وهو مايسمى **(باليرقان الولادي)** او ابو صفار ويظهر عادة في اليوم الثاني او الثالث بعد الولادة ويحدث نتيجة لعدة اسباب منها ارتفاع قيم الخضاب عند الطفل وعند محاولة جسم الطفل التخلص من كمية الكريات الحمراء الزائدة ينجم عن تحطمها مادة **البيليروبين** المسؤولة عن حدوث اليرقان واذا ماحدث ارتفاع شديد في قيم البيليروبين يمكن ان يكون اليرقان خطيراً جداً ويؤدي الى تراكم هذه المادة في دماغ الطفل واصابة الطفل بحالة خطيرة تسمى اليرقان النووي وقد تؤدي الى التخلف العقلي والشلل الحركي.

لذا ارتأينا استخدام الاسلوب الاحصائي **التحليل المميز**"discriminate analysis" بأعتباره أحد المواضيع المهمة في التحليل المتعدد المتغيرات والذي يسهم في كيفية التمييز بين مجموعتين أو أكثر من المتغيرات .

التحليل المميز

إن الفكرة الأساسية من التمييز Discriminate هو التفرقة بين مجتمعات متداخلة أو متشابكة لها الخصائص نفسها أو الصفات ، بمعنى آخر لنفرض إن لدينا مجتمعين أو أكثر ولدينا عينة تحتوي على مجوعة من المشاهدات من كل مجتمع فإن وظيفة التحليل المميز هي إيجاد دالة يمكن بواسطتها تصنيف أو تمييز المشاهدات الجديدة إلى مجتمعاتها الأصلية، لقد كانت أول فكرة تمييز بين مجتمعات مختلفة على أساس التشابه بين الأشياء للباحث "Karl person" عام 1921 وذلك من خلال قياس المسافة الإحصائية بين عينتين والذي يدعى معامل التشابه النسبي Coefficient Racial likeness ويكتب اختصاراً C.R.L وحسب الصيغة التالية: -

$$\text{C.R.L} = \frac{1}{p}\left[\sum_{i=1}^{p}\left(\frac{di^2}{\frac{S_{1i}^2}{n_1}+\frac{S_{2i}^2}{n_2}}\right)\right]$$

حيث إن

S_{1i}^2 S_{2i}^2 التباين المقدر للمتغير (i) المناسب في المجموعة الأولى والثانية على التوالي

i=1,2,3,.....p يمثل عدد الصفات المستخدمة

وأن n_2, n_1 حجم العينة الأولى والثانية على التوالي

$$d_i = \overline{X}_{1i} - \overline{X}_{2i}$$

اذ أن

$\overline{X}_{i1}, \overline{X}_{2i}$ تمثل متوسطي العينتين للمتغير i المقاس في المجموعة الأولى والثانية على التوالي.

وأستخدم الباحث " P.C "Mahalanobis مقياساً حديثاً لقياس المسافات بين مراكز مجتمعين وقد سمّى هذا المقياس بمقياس (D^2) وتدعى مسافة مهلنوبيس وصيغتها بالشكل التالي:

$$D^2 = \frac{1}{P}\sum_{i=1}^{P}\left(\frac{\overline{X}_{1i} - \overline{X}_{2i}}{\sigma i}\right)$$

حيث أن :

σ_i تمثل الإنحراف لمجموعة i من الصفات ويمكن كتابتها بشكل مصفوفات وكالآتي:

$$D^2 = (\overline{X}_{1i} - \overline{X}_{2i})' S^{-1} (\overline{X}_{1i} - \overline{X}_{2i})$$

اذ أن

$\overline{X}_{1i}, \overline{X}_{2i}$ هي مصفوفة الفرق بين متوسطي العينتين المأخوذتين من المجموعة (1i ,2i) وأن S^{-1} تمثل معكوس مصفوفة التباين المشترك المقدر لمصفوفة المعالم المجهولة .

وأن ($\overline{X}_{1i}, \overline{X}_{2i}$) S^{-1} هي معاملات الدالة المميزة .

وأوجد الباحث هارولد هوتلنك (Harold Hotelling) عام 1931 مؤشر إحصائي أسماه T^2 والذي يعالج مشكلة إتخاذ قرار حول العينتين اللتين يكون توزيع كل منهما توزيعاً طبيعياً متعدد المتغيرات ولها نفس مصفوفة التباين المشترك وصيغة هذا المؤشر هي :-

$$T^2 = \sum_{i=1}^{P} \sum_{j=1}^{P} \left(\frac{1}{\frac{1}{n_1} + \frac{1}{n_2}} \right) * Sij^{-1}.di.dj$$

وبدرجة حرية n_1+n_2 -2

$$di = \overline{X}_{1i} - \overline{X}_{2i}$$

$$dj = \overline{X}_{1j} - \overline{X}_{2j}$$

وفي عام 1936 قام الباحث فشر Fisher باقتراح الدالة المميزة الخطية للوصول إلى أعلى مجال للفصل بين مجتمعين تابعين إلى نفس المصدر وذلك باستعمال عدد من المتغيرات المرتبطة بتصنيف المفردات حسب مجتمعاتها وهذه أول عملية تمييز حقيقية لتصنيف المفردة إلى أحد المجتمعين وعرف الدالة التمييزة الخطية بالشكل التالي:-

$$Z=\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots\ldots\ldots + \alpha_p x_p$$

وقد أقترح الباحث (ويلد Wuld) عام 1944 بأن تقدير المعالم لدالة فشر أعلاه وتكون باستخدام طريقة الإمكان الأعظم M.L.E وتوالت بعدها البحوث في استخدام الدالة التمييزية وذلك من خلال استخدامها تطبيقياً و نظرياً حيث طورت الطريقة و استحدثت دوال جديدة مثل الدالة التربيعية و اللوجستية .

وإضافة إلى ذلك أهتم الكثير من الباحثين بهذه الطريقة لتكون المعينة لهم في تصنيف وتحديد المشاهدات الجديدة ومجتمعاتها .

هدف البحث

يهدف البحث إلى تحديد أهمية العوامل المؤثرة على اليرقان في الأطفال حديثي الولادة التي تم تحديدها من قبل أطباء أخصائين بامراض الاطفال و باستخدام أسلوب التحليل المميز لغرض معرفة أسباب انتشار هذا المرض .

أنواع الدوال التمييزية

هناك عدة أنواع من الدوال التمييزية والتي يمكن استخدامها بالاعتماد على البيانات المدروسة وخصائصها ومن هذه الدوال :-

١. الدالة المميزة الخطية .
٢. الدالة المميزة التربيعية .
٣. الدالة المميزة اللوجستية .

وسوف يتم التطرق إلى الدالة المميزة الخطية بشيء من التفصيل كونها الدالة المناسبة لنوع وخصائص البيانات في الحالة المدروسة .

الدالة المميزة الخطية (Linear Discriminate Function)

تستخدم الدالة المميزة الخطية عندما تكون المجتمعات المدروسة ذات توزيع طبيعي متعدد المتغيرات بمتجهات متوسطة مختلفة . مصفوفة تباين وتباين مشترك متساوية .

أن عملية التمييز تتلخص بالخطوات التالية :-

١- إيجاد متوسط كل متغير في كل مجموعة ثم إيجاد الفرق بين متوسطي كل متغير في المجموعتين حيث أن متوسط المجموعة الأولى

$$\overline{X}_{(1)} = \sum_{j=1}^{n1} Xj / n_1$$

متوسط المجموع الثانية

$$\overline{X}_{(2)} = \sum_{j=1}^{n2} Xj / n_2$$

أما الفرق بين متوسطي المجموعتين d_i حيث أن $di = \overline{X}(1) - \overline{X}(2)$

٢- إيجاد مجموع المربعات لكل متغير في كل مجموعة وكذلك مجموع حاصل ضرب كل متغيرين داخل كل مجموعة أيضاً

$$Sii = \sum Xi^2 - (\sum \overline{Xi})^2 / n$$

$$Sij = \sum Xi.Xj \ - (\sum Xi)(\sum Xj) \ / n$$

٣- إيجاد التباين والتباين المشترك المدمج (داخل المجاميع)

$$Vii = \frac{Sii_{(1)} + Sii_{(2)}}{n_1 + n_2 - 2}$$

$$Vij = \frac{Sij_{(1)} + Sij_{(2)}}{n_1 + n_2 - 2}$$

٤- دالة التمييز تكتب كالآتي

$$y = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \ldots\ldots\ldots + \alpha_r x_r$$

حيث $\alpha_1, \alpha_2, \alpha_3, \ldots\ldots, \alpha_r$ تختار بحيث تعطي أعلى تمييز بين المجموعتين والمقصود بأعلى تمييز بين المجموعتين هو أن الاختلاف بين قيم المجموعتين Between- Group variation أكبر بكثير من الاختلافات بين قيم داخل

المجموعتين Within – Group Variation فإذا رمزنا لنسبة الاختلافات بين المجموعتين إلى الاختلافات داخل المجموعتين بالرمز λ.

$$\lambda = \frac{between - group.\text{var}iation}{withen - group.\text{var}iation}$$

فإننا نختار $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_r$ بحيث تكون λ أكبر مما يمكن هذا ويمكن حساب أهمية كل متغير .

$$\mathbf{X_i^* = \alpha} \sqrt{Vii}$$

حيث ان

$$\alpha = s^{-1}(\overline{X}_1 - \overline{X}_2) \quad \text{و} \quad \alpha' = (\overline{X}_1 - \overline{X}_2) \quad s^{-1}$$

٥- إيجاد نقطة الفصل

لتصنيف الأفراد فإننا نحتاج إلى نقطة تفصل بين المجموعتين بحيث إذا قلت (مجموعة نقاط الأفراد) عن تلك النقطة فإن الفرد عندئذ يتبع إلى مجموعة معينة وإذا زادت مجموع نقاطه عن تلك النقطة الفاصلة أصبح تابعاً للمجموعة الأخرى .

$$Z = \frac{\overline{Y}_1 + \overline{Y}_2}{2}$$

حيث أن :-

$$\overline{y}_1 = \overline{x}(1).S^{-1}(\overline{x}(1) - \overline{x}(2))$$
$$\overline{y}_2 = \overline{x}(2).S^{-1}(\overline{x}(1) - \overline{x}(2))$$

٦- قاعدة التصنيف

يمكن وضع قاعدة التصنيف وكالاتي:

المشاهدة x تعود للمجتمع الأول إذا كان $y \geq Z$

المشاهدة x تعود للمجتمع الثاني إذا كان $y < Z$

٧- خطأ التصنيف The Probability of Misclassification

هو احتمال تصنيف مشاهدة معينة إلى المجموعة الأولى بينما هي تعود في الحقيقة إلى المجموعة الثانية وبالعكس نفترض لحساب خطأ التصنيف إن حجم العينة يكون كبير لذلك توزيع المشاهدات يقترب من التوزيع الطبيعي (حسب نظرية الحد المركزي) حيث يعتمد على أن العينة تتوزع توزيعا طبيعيا أو يقترب من التوزيع الطبيعي.

$$P_{12} = \Phi(-1/2D)$$

حيث أن

$$D^2 = (\overline{X}_{(1)} - \overline{X}_{(2)})' * S^{-1}(\overline{X}_{(1)} - \overline{X}_{(2)})$$

ويتم إيجاد قيمة P_{12} من جداول التوزيع الطبيعي القياسي . أن خطأ التصنيف هو عامل مهم لإثبات كفاءة الدالة المميزة والدالة التي تعطي أقل خطأ تصنيف هي الدالة الأكثر كفاءة وتكون الأفضل من بين دوال التميز.

٨- إختبار كفاءة الدالة المميزة الخطية Testing of Linear Discriminate Function

يتم التأكد من معنوية الدالة المميزة الخطية عن طريق اختبار الفرضية .

H0: $\alpha_1= \alpha_2= \alpha_3=\ldots= \alpha_r$

ويتم الاختبار باستخدام اختبار F الذي يعتمد على إحصاءه هوتلنك (Hotelink) T^2 .

$$F = \left[\frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p}\right] * T^2$$

$$T^2 = \frac{n_1.n_2}{n_1 + n_2} D^2$$

حيث أن

p : عدد المتغيرات

D^2 : تعني إحصاءه مهلنوبيس

وعند مقارنتها بقيمة F الجدولية فإذا كانت المحسوبة أكبر من F الجدولية بدرجة حرية (n_1+n_2-p-1) هذا يعني أن الاختلافات معنوية بين المجموعتين وأن دالة التمييز لها فائدة كبيرة في التمييز أو تصنيف الأفراد.

جمع البيانات Data gathering

إن إسلوب جمع البيانات هو عبارة عن الإستماره الإحصائية التي أعدت لهذا الغرض والتي تضمنت دراسة حول مرض اليرقان في محافظة القادسية وقد شملت معظم الأقضية والنواحي في المحافظة .

وشملت الدراسة 80 مشاهدة من الذكور والإناث المصابين بمرض اليرقان وقد تم تقسيم المصابين إلى مجموعتين .

١- المجموعة الأولى: وقد شملت الأطفال الذين تم إجراء عملية تبديل الدم لهم ويرمز لها بالرمز (٠) .

٢- المجموعة الثانية : وقد شملت الأطفال الذين لم يتم إجراء عملية تبديل الدم لهم ويرمز لها بالرمز (١) .

حيث كانت المتغيرات x_i ، i=1,2,3,……..,8 هي كما موصوفة في التالي:

X_1 يمثل نوع الجنس للطفل حيث
- 0 يرمز للذكور
- 1 يرمز للإناث

X_2 يمثل العمر للطفل

X_3 يمثل الوزن للطفل

X_4 يمثل نسبة TSP (Total serum bilirubin) اذا كان الطفل مصاب سوف يحتوي مصل الدم على نسبة البيليروبين وبنسب معينة وحسب شدة الحالة.

X5 يمثل صنف الدم للمولود

حيث انه
- يرمز لـ A بالرمز 1
- يرمز لـ B بالرمز 2
- يرمز لـ AB بالرمز3
- يرمز لـ O بالرمز 4

X_6 يمثل نوع صنف الدم RH (Rhesus factor) وهي عبارة عن وجود او عدم وجود نوع من البروتين على سطح خلايا كريات الدم الحمراء ففي حالة وجوده تكون الفصيلة موجبة وعند عدم وجوده تكون الفصيلة سالبة.

حيث انه
- يرمز لـ RH^+ بالرمز 1
- يرمز لـ RH^- بالرمز 0

X_7 يمثل صنف الدم للأم

حيث انه
- يرمز لـ A بالرمز 1
- يرمز لـ B بالرمز 2
- يرمز لـ AB بالرمز3
- يرمز لـ O بالرمز 4

X_8 يمثل نوع صنف الدم للأم
- يرمز لـ RH^+ بالرمز 1
- يرمز لـ RH^- بالرمز 0

التحقق من شروط الدالة التمييزية

١- البيانات تتوزع توزيع طبيعي.

لقد تم استخدام اختبار كولمكروف – سميرنوف Kolmgrov-Smirnov لاختبار فرضية العدم التي تفترض بان البيانات تتوزع توزيعا طبيعيا بمتوسط وتباين معينين وقد تم اختبار المتغيرات كل على حده وذلك بالاعتماد على الخاصية التي تقول بانه اذا كانت هناك عدة متغيرات تتوزع توزيعا طبيعيا فان مجموع هذه المتغيرات يتوزع توزيعا طبيعيا ايضا ، حيث كانت قيم اختبار كولمكروف–سميرنوف وقيم المتوسط والانحراف المعياري لكل متغير كما موضحة في الجدول التالي:

المتغير	قيمة الاحصاءة D	المتوسط	الانحراف المعياري
X1	٠.٠٩٧	٠.٥٥	٠.٥٠١
X2	٠.١٢٦	٦.١٦٣	٣.٥١٦
X3	٠.٠٥٢	٢.٧٩٩	٠.٦٩٤
X4	٠.٠٣٩	٢٠.٤١٦	٤.٢١٦
X5	٠.٠٩٨	٢.٤١٣	١.٢٥
X6	٠.٠٩٥	٠.٨٢٥	٠.٣٨٢
X7	٠.٠٨٩	٢.٦٦٣	١.٢٥٢
X8	٠.٠٩٨	٠.٧٥	٠.٤٣٦

وبعد مقارنة قيم D المحسوبة للمتغيرات الثمانية مع القيمة $1.36/n^{1/2}$ والتي تساوي ٠.١٥٢ حيث ان حجم العينة n مساوية الى ٨٠ ، ومن خلال الجدول اعلاه نلاحظ ان جميع قيم D هي اقل من القيمة الجدولية والتي تقودنا الى قبول فرضية العدم وذلك بان البيانات تتوزع توزيعا طبيعيا.

٢- مصفوفة التباين والتباين المشترك متساوية للمجتمعين والبيانات تم أخذها من مجتمع طبيعي ومن ثم سوف نقوم بإجراء اختبار لإثبات تساوي مصفوفة التباين والتباين المشترك للمجتمعين حيث ان فرضية العدم المطلوب اختبارها هي :

$$H_0 = \sum 1 = \sum 2$$
$$H_1 = \sum 1 \neq \sum 2$$

حيث ان احصاءة الاختبار والتي تتبع توزيع مربع كاي وكما موضح بالتالي

$$mc^{-1} \approx X^2(k-1)(p-1)$$

P :عدد المتغيرات

K :عدد المجاميع

وان

$$m = \sum ni * Ln|S| - \sum ni * Ln|Si|$$

$$s = \begin{bmatrix} .25 & .1 & .08 & -.0 & .05 & -.02 & -.01 & .01 \\ .11 & 12.4 & .79 & 1.5 & .39 & -.40 & -.39 & -.15 \\ .08 & .8 & .48 & .8 & .07 & -.06 & .06 & .01 \\ -.04 & 1.5 & .81 & 19.5 & -.02 & -.36 & 1.79 & -.04 \\ .05 & .4 & .07 & -.0 & 1.56 & .01 & .61 & -.06 \\ -.02 & -.4 & -.06 & -.4 & .01 & .15 & .05 & .06 \\ -.01 & -.4 & .06 & 1.8 & .61 & .05 & 1.57 & -.07 \\ .01 & -.1 & .01 & -.0 & -.06 & .06 & -.07 & .19 \end{bmatrix}$$

حيث ظهرت قيمة m وكانت مساوية الى

$$m = 5.617$$

$$C^{-1} = 1 - \frac{2P^2 + 3P - 1}{6(K-1)(P+1)} \left[\sum \frac{1}{ni} - \frac{1}{\sum ni} \right]$$

$C^{-1}=0.8951$

$m.C^{-1}=5.617* 0.8951=5.028$

ثم تم حساب قيمة X^2 الجدولية وكانت $X^2_{7,0.05}=14.067$

إذن القيمة المحتسبة أقل من القيمة الجدولية

إذن تقبل فرضية H_0 أي أن $\sum_1 = \sum_2$

نلاحظ بأن شرطي الدالة التمييزية الخطية قد تحقق

ويتم إيجاد دالة التمييز الخطية وحسب الصيغة التالية :

$$Y = X' * S^{-1}(\overline{X}_1 - \overline{X}_2)$$

$$\therefore Y = 0.255373X_1 + 0.273871X_2 - 0.26115X_3 - 0.76286X_4 + 0.233727X_5 + 0.189772X_6 + 0.244201X_7 + 0.148457X_8$$

ثم استخرجنا قيم $\bar{y}_1, \bar{y}_2$ وكانت (-4.752,-6.994) على التوالي

وبتطبيق المعادلة التالية

$$Z = \frac{\overline{Y}_1 + \overline{Y}_2}{2}$$

حصلنا على قيمة النقطة الفاصلة وكانت مساوية الى (-5.873)

<u>قاعدة التصنيف</u>

المشاهدة x تعود للمجموعة الأولى اذا كانت $y\text{-}z > 0$

المشاهدة x تعود للمجموعة الثانية اذا كانت $y - z \leq 0$

٤- أهمية كل متغير

يمكن الحصول على أهمية كل متغير باستخدام القانون التالي :

$$X_i^* = \alpha\sqrt{v_{ii}}$$

وكانت القيم كالاتي

$X_1^* = 0.255373\sqrt{0.253} = 0.1284$

$X_2^* = 0.273871\sqrt{12.328} = 0.9616$

$X_3^* = 0.26115\sqrt{0.491} = 0.1282$

$X_4^* = 0.76286\sqrt{19.5} = 3.369$

$X_5^* = 0.233727\sqrt{1.558} = 0.2917$

$X_6^* = 0.189772\sqrt{0.146} = 0.0725$

$X_7^* = 0.244201\sqrt{1.542} = 0.3032$

$X_8^* = 0.148457\sqrt{0.192} = 0.0651$

وبعد حساب أهمية كل متغير نجد ترتيبها ويكون كالآتي حسب درجة الأهمية

درجة أهمية كل متغير	المتغير	وصف المتغير
1	X_4^*	نسبة TSP
2	X_2^*	عمر الطفل
3	X_7^*	صنف دم الام
4	X_5^*	صنف دم الطفل
5	X_1^*	الجنس
6	X_3^*	وزن الطفل
7	X_6^*	Rh لدم الطفل
8	X_8^*	Rh لدم الام

إيجاد خطأ التصنيف

بعد حساب أهمية كل متغير نجد خطأ التصنيف حسب القانون التالي:-

$$D^2 = (\overline{X}_{(1)} - \overline{X}_{(2)})' * S^{-1}(\overline{X}_{(1)} - \overline{X}_{(2)})$$

$D^2 = 2.2420$

$D = 1.4973$

$P_{12} = P_{21} = \Phi(-1/2\ D)$

$= \Phi(-1/2 * 1.4973)$

$= \Phi(-0.7487)$

ومن خلال جداول التوزيع الطبيعي القياسي يتم استخراج القيمة وكانت مساوية الى

$P_{12} = (\ 0.2249\ )$

ونلاحظ من خلال قيمة خطا التصنيف بان قيمتها قليلة مما يدل على قوة الدالة التمييزية على تصنيف المشاهدات الى مجتمعاتها الاصلية

اختبار كفاءة الدالة المميزة الخطية

لاختبار كفاءة الدالة المميزة الخطية نختبر الفرضية التالية :

$H_0 : M_1 = M_2$

$H_1 : M_1 \neq M_2$

ثم تحسب قيمة T^2 حسب الصيغة التالية :

$$T^2 = \frac{n_1 . n_2}{n_1 + n_2} D^2$$

$T^2 = 44.84$

وبعد حساب قيمة T^2 تحسب قيمة F حسب القانون التالي:

$$F = \left[\frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)p} \right] * T^2$$

F = ٥.587

وبعد حساب قيمة F الجدولية نجد أن قيمة F المحسوبة أكبر من قيمة F الجدولية إذن نرفض الفرضية H_0 وبالتالي فإن دالة التمييز الخطية قادرة على تصنيف كل مشاهده إلى مجتمعها الأصلي والمشاهدات ذات كفاءة عالية أي أن الاختلافات بين المجتمعين معنوية .

الاستنتاجات

نستنتج بأن نسبة TSB لها الأثر الكبير في إجراء عملية تبديل الدم للطفل المريض.

ويأتي تاثير المتغيرات عمر الطفل ، صنف الدم للأم ، صنف الدم للمولود ، الجنس ، وزن الطفل ، Rh لدم الطفل ، Rh لدم الام على التوالي.

المصادر

١- الراوي خاشع محمود "المدخل إلى تحليل الإنحدار" / كلية الزراعة والغابات جامعة الموصل .

2- Cooley , W.W. cohnes , P.v. (1971), " Multivariate Data Analysis " Jonnwiley and sons .Newyork.

3- Anderson , T.w ((1958)) An Introduction to Multivariate statistical Analysis Jonn wiley and sons, Inc.

4- Krishnaiah , P,R ((1977)) Multivariate Analysis "North – Holland Publishing company-New York – Oxford

5- Krzanowski, W.J. (1995). "Multivariate Analysis Classification, Covariance structures", John wiley

* Stons, Inc, New York.