

التحليل التمييزي وفعاليته في تصنيف تأثير وزن الحقيبة المدرسية على صحة التلاميذ "دراسة تطبيقية على تلاميذ مرحلة التعليم الأساسي بمدينة الزاوية"

عبدالحليم مولود الصويحي¹

قسم الإحصاء، كلية العلوم

جامعة الزاوية، ليبيا

halimsuaiee@zu.edu.ly

فاطمة خليفة بنيني²

قسم الإحصاء، كلية العلوم

جامعة الزاوية، ليبيا

f.banini@zu.edu.ly

ملخص Abstract

تم استخدام تحليل التمييز لتصنيف تأثير وزن الحقيبة المدرسية على آلام الظهر لمجموعة من تلاميذ التعليم الأساسي في المدارس الحكومية بمدينة الزاوية إلى مجموعتين اعتماداً على متغيرات ذات صفات تمييزية. وباستخدام دالة التمييز الخطية تم تصنيف تأثير وزن الحقيبة على أساس مجموعة من المتغيرات المصاحبة للآلام الظهر، حيث كانت نسبة التصنيف الصحيح الكلي لها 63.3% في حين أن نسبة التصنيف الخاطئ كان 36.7% وأن أكثر المتغيرات أهمية في التصنيف هو متغير وزن الحقيبة إلى وزن الجسم ثم يليه متغير النوع ثم متغير العمر.

مقدمة: Introduction

يعد تحليل التمييز Discriminant Analysis أحد طرق تحليل متعدد المتغيرات Multivariate Analysis الذي يهتم بدراسة تأثير مجموعة من العوامل في مفردات مجموعات مختلفة متميزة عن بعضها البعض. ويستخدم للتمييز بين مجموعتين أو أكثر تتشابه في الكثير من الصفات على أساس وجود عدد من المتغيرات. ويستفاد من تحليل التمييز في تصنيف مفردة واحدة أو أكثر إلى مجتمع أو مجموعة من المجتمعات بالاعتماد على متغيرات لها صفات معينة. كذلك يستفاد منه في التعرف على المتغيرات التي تساهم في عملية التصنيف. يستخدم التحليل التمييزي في مجالات مختلفة، وكمثال على ذلك يستخدم في حالة تصنيف المرضى أو المصابين بمرض معين بالاعتماد على مجموعة من المعلومات عن حالتهم الصحية.

مشكلة البحث

إن الاهتمام بتلاميذ المرحلة من التعليم الأساسي يعد من التحديات التي تواجه المسؤولين من كل الجوانب وخاصة الصحية منها، حيث تعتبر من الأولويات توفير بيئة صحية للتلميذ وعدم تحميله أكثر من طاقته. وتعد وزن حقيبة التلميذ في مراحله التعليمية الأولى من الأمور المهمة والمؤثرة في صحة التلميذ، حيث أن الوزن الزائد لحقيبة التلميذ يؤدي إلى حدوث آلام مختلفة للتلميذ وخاصة آلام الظهر. وبالتالي يمكن تحديد مشكلة البحث في إمكانية تصنيف التلاميذ بناء على شعورهم بالآلام الظهر استناداً إلى وزن الحقيبة المدرسية والعوامل المؤثرة فيها.

الهدف من البحث

يهدف البحث إلى تشخيص العوامل المؤثرة في وزن الحقيبة المدرسية على آلام الظهر لمجموعة من تلاميذ التعليم الأساسي ومعرفة تأثير كل عامل من هذه العوامل من خلال استخدام الدالة التمييزية الخطية.

مفهوم التحليل التمييزي

يعتبر التحليل التمييزي أحد أهم الأساليب الإحصائية لمتعدد

الكلمات المفتاحية: التحليل التمييزي، وزن الحقيبة المدرسية، تلاميذ التعليم الأساسي

إلى مجموعات وهي المتغير الأصلي مع مجموعة من المتغيرات المستقلة. حيث يفضل استخدام التحليل التمييزي عندما يكون المتغير التابع وصفي وله صفتين أو أكثر، بالإضافة إلى وجود مجتمعين أو أكثر. وتكون هذه المجتمعات متشابهة ولكن منفصلة احصائياً ومنها تكوين قاعدة للفصل بين هذه المجتمعات تستخدم في تصنيف مشاهدات جديدة غير معروف المجتمع الذي تنتمي إليه (العباسي، 2011).

أهداف التحليل التمييزي

1. تصميم التركيبات الخطية للمتغيرات الأفضل في التمييز بين فئات المتغير التابع.
2. التحقق فيما إذا كان هناك فروق ذات دلالة احصائية بين المجموعات فيما يتعلق بالمتغيرات.
3. تصنيف المتغيرات التي تسهم بأكبر قدر من الاختلاف بين فئات المتغير التابع.
4. تقييم دقة التصنيف كنسبة مئوية (جودة، 2008).

مسلمات التحليل التمييزي

يستند التحليل التمييزي إلى عدد من المسلمات نوجزها كما يلي:

1. أن تكون المتغيرات التابعة الكمية لها توزيع طبيعي، لهذا يفضل أن تكون العينات الخاصة بالدراسة بحجم كبير للحصول على نتائج أكثر مصداقية.
2. أن تكون تباينات وتغايرات المتغيرات التابعة في المجتمع متجانسة في جميع مستويات العامل، حيث يتم التأكد من ذلك باستخدام اختبار Box's M (نجيب و الرفاعي، 2006).
3. أن يتم اختيار العينة بشكل عشوائي، كما يجب أن تكون درجة أي مشاهدة في العينة في أي متغير مستقل عن جميع مشاهدات العينة الآخرين. في حالة عدم تحقق هذا الشرط، فإن اختبار الدلالة للتحليل التمييزي يصبح غير موثوق (أبو علام، 2003).

أنواع التحليل التمييزي

توجد ثلاثة أنواع متاحة للتحليل التمييزي وذلك حسب طريقة ادخال

المتغيرات إلى التحليل، وهي كالتالي:

1. التحليل التمييزي المباشر Direct Discriminant Analysis (DDA)

حيث يتم ادخال جميع المتغيرات المستقلة إلى التحليل مرة واحدة وبدون استثناء ودون اعطاء أي أهمية لترتيب دخولها.

2. التحليل التمييزي الهرمي Hierarchical Discriminant analysis (HAD)

حيث يتم ادخال المتغيرات طبقاً لجدول ينظمه الباحث من حيث أهمية المتغيرات المستقلة والترتيب الذي يعتقد أنه مناسب.

التحليل التمييزي التدريجي Stepwise Discriminant Analysis (SDA)

حيث يتم تحديد معيار احصائي يحدد أولوية ادخال المتغيرات إلى التحليل (نجيب و الرفاعي، 2006). وتعتبر هذه الطريقة الأكثر استخداماً إذ يكون فيها ترتيب إضافة المتغيرات المستقلة إلى التحليل واستبعادها منه تبعاً لمعايير احصائية.

شروط استخدام التحليل التمييزي

بعد تحديد المتغير التابع والمتغيرات المستقلة لابد من التأكد من تحقق الشروط اللازمة لإجراء التحليل، يمكن تلخيصها في النقاط التالية:

وكما كانت هناك فروق معنوية بين المجموعات، فإن ذلك يدل على أن الدالة التمييزية الخطية قابلة للتمييز بدرجة عالية.

4. إيجاد قيم وليكس لامدا (Wilks Lambda)

يتم حساب احصاءة وليكس لامدا من خلال الصيغة التالية:

$$\Lambda = \frac{|W|}{|T|}$$

حيث أن:

W : تمثل مصفوفة التباين والتغاير داخل المجموعات

T : تمثل مصفوفة التباين والتغاير الكلي للمجموعات

وتبين احصاءة وليكس لامدا درجة التباعد بين المجموعات وتتراوح قيمته بين الصفر والواحد. فكلما كانت قيمته قريبة من الصفر، يشير ذلك لوجود تفرقة عالية بين المجموعات ويدل ذلك على قوة التمييز. أما إذا كانت القيمة قريبة من الواحد، فإن ذلك يشير إلى عدم وجود تمييز بين المجموعات. حيث تتكون الدالة من المتغيرات التي لها أقل قيمة لإحصاءة وليكس لامدا (الحسن و رحمة، 2017).

إيجاد نقطة الفصل Cut Point

لتصنيف المشاهدات نحتاج إلى نقطة تفصل بين المجموعتين بحيث إذا كانت مجموعة نقاط المشاهدات أقل من نقطة الفصل، فإن المشاهدات تصنف ضمن مجموعة معينة. أما إذا زاد مجموع نقاط المشاهدات عن نقطة الفصل، فإن المشاهدات تصنف ضمن المجموعة الأخرى. وتحسب نقطة الفصل من الصيغة التالية:

$$y_c = \frac{\bar{y}_1 + \bar{y}_2}{2}$$

أي أنه إذا كانت $y_i - y_c > 0$ فإن المشاهدات تصنف ضمن المجموعة الأولى.

أما إذا كانت $y_i - y_c \leq 0$ فإن المشاهدات تصنف ضمن المجموعة الثانية.

5. احتمال خطأ التصنيف Probability of Misclassification

ويقصد به احتمال تصنيف مشاهدة معينة إلى المجموعة الأولى وهي في الحقيقة تنتمي إلى المجموعة الثانية وبالعكس. ويتم حساب احتمال خطأ التصنيف من خلال الصيغة التالية:

$$Y = b'S$$

حيث أن:

$$b' = (\bar{X}_1 - \bar{X}_2)'S^{-1}$$

Φ : تمثل دالة التوزيع الطبيعي المعياري

D : تمثل مسافة مهنوبس

بعد حساب القيمة داخل القوسين، يتم استخراج الاحتمال المقابل لها من جدول التوزيع الطبيعي المعياري. فكلما كان الاحتمال صغير واقترب من الصفر دل ذلك على أن احتمال التصنيف الخاطئ ضئيل، وبالتالي قوة الدالة في عملية التصنيف. أما إذا كانت قيمة الاحتمال تقترب من الواحد، فإن ذلك يدل على أن احتمال التصنيف الخاطئ كبير وبالتالي ضعف الدالة في عملية التصنيف.

تقييم أداء دالة التصنيف

تتمثل إحدى الطرق المهمة للحكم على أداء أي إجراء تصنيف في حساب "معدلات الخطأ". هناك مقياس يسمى معدل الخطأ الظاهري (APER)، يمكن حسابه بسهولة من مصفوفة التصنيفات، والتي تظهر عناصر المجموعة الفعلية مقابل المجموعة المتوقعة (Johnson و

$$D^2 = (\bar{X}_1 - \bar{X}_2)'S^{-1}(\bar{X}_1 - \bar{X}_2)$$

حيث أن

n_1m : عدد العناصر من المجموعة الأولى المصنفة بشكل خاطئ في المجموعة الثانية.

$$F = \frac{(n_1 + n_2 - p - 1)T^2}{(n_1 + n_2 - 2)p}$$

n_2m : عدد العناصر من المجموعة الثانية المصنفة بشكل خاطئ من المجموعة الأولى.

الجانب التطبيقي

1. يجب أن يكون اختيار العينة بشكل عشوائي وحجمها كبير بحيث يسمح بإجراء كل التحليلات المصاحبة لعملية التحليل.
2. أن تكون المجتمعات الإحصائية موضوع الدراسة تتوزع توزيعاً طبيعياً.
3. تجانس مصفوفة التباين والتغاير Variance Covariance Matrix في المجتمعات الإحصائية موضوع الدراسة (Alvin، 2002).
4. عدم وجود ارتباط قوي بين المتغيرات المستقلة.

الدالة التمييزية الخطية: The Linear Discriminant Function

هناك حالتين للتمييز بين المجموعات وهي إما التمييز بين مجموعتين أو التمييز بين أكثر من مجموعتين. هذه الدراسة ركزت على التمييز بين مجموعتين كون أن مشاهدات المجتمع قسمت إلى مجموعتين فقط.

الدالة التمييزية الخطية لمجموعتين The Linear Discriminant Function for Two Groups

لو فرضنا أن لدينا مجموعتين عشوائيتين من مجتمعين طبيعيين الأولى بحجم n_1 ولها p من المتغيرات بمتوسط حسابي \bar{X}_1 ومصفوفة تباين مشترك S . والثانية بحجم n_2 ولها p من المتغيرات المستقلة بمتوسط حسابي \bar{X}_2 ومصفوفة تباين مشترك S . بذلك يمكن صياغة تركيب خطي يشمل المشاهدات من المجموعتين الغرض منه التمييز بين المجموعتين (الجاعوني و عدنان، 2007).

إن الدالة التمييزية الخطية عبارة عن نموذج يمكن صياغته اعتماداً على مؤشرات العينة التي تم اختيار مشاهداتها ووضعت في مجموعتين مختلفتين، وبواسطة هذه الدالة نستطيع أن نختبر المشاهدة ونحدد انتمائها إلى أي مجموعة. عليه يمكن صياغة الدالة التمييزية الخطية على الشكل التالي:

$$Y_i = b_1X_1 + b_2X_2 + \dots + b_pX_p, i = 1, 2, \dots, r$$

حيث أن:

Y : القيمة التمييزية المعيارية

p : عدد المتغيرات المستقلة الداخلة في الدالة

b : معاملات الدالة التمييزية المعيارية

r : عدد الدوال التمييزية

وتكتب الدالة التمييزية الخطية باستخدام المصفوفات بالشكل التالي:

حيث أن:

و عليه تكون الدالة التمييزية الخطية على النحو التالي:

من أجل إيجاد معادلة التمييز أعلاه، يتم اتباع الخطوات التالية:

1. تقدير معاملات الدالة التمييزية

يتم تقدير المعلمات للدالة التمييزية من خلال الصيغة التالية:

حيث أن:

\bar{X}_1 ، \bar{X}_2 متجهها متوسطي المجموعتين الأولى والثانية

S : مصفوفة التباين والتغاير المشترك

2. إيجاد مربع المسافة Square Distance

والتي تسمى مسافة مهنوبس Mahalanobis Distance

وتحسب من الصيغة التالية:

3. اختيار المتغيرات المستقلة التي APER منها الدالة التمييزية الخطية من خلال الصيغة التالية:

وبدرجة حرية $(p, n_1 + n_2 - p - 1)$

$$T^2 = \frac{n_1n_2}{n_1+n_2} D^2$$

حيث أن: D^2 مسافة مهنوبس

حيث يتم اختيار المتغيرات التي لها أعلى قيمة F ، وتمثل معدل (F) مساهمة المتغيرات المستقلة في التمييز بين المجاميع (يعقوب، 2017).

وصف عينة البحث

تم جمع البيانات من المدارس الحكومية للتعليم الأساسي بمدينة الزاوية. 800 تلميذ تم اختيارهم بشكل عشوائي خلال الفترة من يناير إلى مارس 2016 لمعرفة تأثير وزن الحقيبة المدرسية على آلام الظهر، حيث قسمت البيانات إلى مجموعتين. المجموعة الأولى تمثل التلاميذ الذين لا يشعرون بالألم بالظهر، والمجموعة الثانية تمثل التلاميذ الذين يشعرون بالألم بالظهر. اعتبر الشعور بالألم من عدمه هو المتغير التابع Dependent Variable حيث أعطيت القيم (1) للتلاميذ الذين لا يشعرون بالألم، والقيمة (2) للتلاميذ الذين يشعرون بالألم. أما المتغيرات المستقلة Independent variable فهي كالتالي:

X_1 : نوع التلميذ حيث (1) تعني ذكر و (2) تعني أنثى

X_2 : عمر التلميذ، X_3 : وزن التلميذ

X_4 : طول التلميذ، X_5 : وزن حقيبة التلميذ

X_6 : نسبة وزن حقيبة التلميذ إلى وزنه

تم استخدام البرنامج الإحصائي SPSS V23 لتحليل البيانات والحصول على النتائج.

اختبار التوزيع الطبيعي للبيانات:

لمعرفة ما إذا كانت البيانات تتبع التوزيع الطبيعي أم لا، تم استخدام اختبار Kolmogorov-Smirnov من خلال البرنامج الإحصائي SPSS V23 حيث كانت النتائج كما في الجدول رقم (1).

جدول رقم (1)

نتائج اختبار Kolmogorov-Smirnov

| المتغير | Kolmogorov-Smirnov Test | | |
|---------------------------|-------------------------|-----|-----------|
| | Sig | N | Statistic |
| النوع | 0.000 | 800 | 0.356 |
| العمر | 0.000 | 800 | 0.229 |
| الوزن | 0.000 | 800 | 0.225 |
| الطول | 0.000 | 800 | 0.382 |
| وزن الحقيبة | 0.000 | 800 | 0.156 |
| وزن الحقيبة إلى وزن الجسم | 0.000 | 800 | 0.310 |

يتبين من نتائج اختبار التوزيع الطبيعي أن متغيرات الدراسة لا تتوزع توزيع طبيعي، ولكن بالنظر لكون أن حجم البيانات كبير (800 مشاهدة) فيمكن اعتبار أن البيانات تتوزع بالتقريب إلى التوزيع الطبيعي وذلك حسب نظرية النهاية المركزية.

التأكد من عدم وجود ارتباط ذاتي بين المتغيرات المستقلة

تم استخدام مقياس Variance Inflation Factor (VIF) لاختبار مشكلة الارتباط الذاتي

جدول رقم (2)

نتائج قيم variance inflation factor (VIF)

| VIF | Statistic |
|-------|---------------------------|
| 1.016 | النوع |
| 1.200 | العمر |
| 1.212 | الوزن |
| 1.016 | الطول |
| 1.562 | وزن الحقيبة |
| 1.465 | وزن الحقيبة إلى وزن الجسم |

يتضح من جدول (2) أن قيم VIF لكل المتغيرات المستقلة أقل من 5 مما يدل على عدم وجود مشكلة الارتباط الذاتي.

اختبار تجانس التباين والتباين المشترك

Homogeneity Test of Variance-Covariance Matrix

لمعرفة مدى تجانس مشاهدات المجموعات، تم استخدام اختبار (Box's M) فكانت النتائج كما في الجدول رقم (3)

جدول (3)

نتائج اختبار Box's M

| Sig | df2 | df1 | F | Box's M |
|-------|-----------|-----|-------|---------|
| 0.000 | 6774234.8 | 21 | 632.5 | 6690.7 |

نلاحظ أن مستوى الدلالة Sig.=0.000 وهي أقل من 0.05 مما يدل على عدم التجانس بين مصفوفات التباينات المشتركة، وهنا عادة يحدث عندما تكون حجم البيانات كبير حيث في هذا البحث كان حجم البيانات 800 مشاهدة حيث هذا الاختبار حساس لاختبار تجانس البيانات، ولكن دالة التمييز تبقى قوية حتى في حالة عدم توفر شرط التجانس في حالة وجود قيم متطرفة. عليه يمكن التعامل مع قيم Log Determinant لتفسير نتائج اختبار تجانس التباين والتباين المشترك حيث كانت قيم Log Determinant متقاربة ومحصورة بين (11.362-8.077)، لذلك يمكن افتراض تجانس مصفوفات التباينات المشتركة [9].

اختبار معنوية المتغيرات الداخلة في الدالة التمييزية

تم اختبار معنوية المتغيرات لمعرفة أهمية كل متغير بشكل منفرد ومدى تأثيره في بناء الدالة التمييزية الخطية، حيث كانت النتائج كما مبينة بالجدول رقم (4)

جدول رقم (4)

نتائج اختبار معنوية الدالة التمييزية

| sig | df2 | df1 | F | Wilk's Lambda | Variable |
|-------|-----|-----|-------|---------------|---------------------------|
| 0.418 | 798 | 1 | 0.66 | 0.999 | النوع |
| 0.004 | 798 | 1 | 8.54 | 0.989 | العمر |
| 0.078 | 798 | 1 | 3.12 | 0.996 | الوزن |
| 0.699 | 798 | 1 | 0.15 | 1.000 | الطول |
| 0.000 | 798 | 1 | 48.67 | 0.942 | وزن الحقيبة |
| 0.000 | 798 | 1 | 38.08 | 0.954 | وزن الحقيبة إلى وزن الجسم |

يتبين من جدول (4) أن متغيرات العمر ووزن الحقيبة ووزن الجسم إلى وزن الجسم هي أهم المتغيرات المستقلة الداخلة في الدالة التمييزية، حيث كانت لها أكبر قيم F وأصغر قيم Wilk's Lambda. وهي متغيرات معنوية من خلال قيم sig. حيث أنها أقل من 0.05. بينما تم استبعاد متغيرات النوع والوزن والطول حيث أن قيم sig لها كانت أكبر من 0.05 وبالتالي فهي غير معنوية.

اختبار معنوية الدالة التمييزية

الجدول التالي يوضح الإحصاءات المتعلقة بدالة التمييز

جدول رقم (5)

الإحصاءات المتعلقة بدالة التمييز

| Test of Function | Wilk's Lambda | Chi-Square | df | Sig |
|------------------|---------------|------------|----|-------|
| 1 | 0.921 | 65.144 | 6 | 0.000 |

يتضح من جدول (5) أن قيمة Wilk's Lambda=0.921 وقيمة Chi-square=65.144 وهي دالة إحصائية حيث أن قيمة Sig.=0.000 وهذا يدل على القدرة الجيدة للدالة على التمييز بين المجموعتين. وبذلك فإن الاختلاف في التصنيف بين المجموعتين هو اختلاف معنوي ولا يعود للصدفة.

تقدير معاملات الدوال التمييزية الخطية

تم تقدير معاملات الدوال التمييزية الخطية للمجموعتين، حيث من خلال هذه المعاملات يمكننا معرفة مدى تأثير المتغيرات على النموذج. إذ كلما كانت قيمة المعامل المطلقة كبيرة، جل ذلك على المساهمة العالية للمتغير. والجدول التالي رقم (6) يبين قيم المعاملات المعيارية للمتغيرات الداخلة في التحليل.

تقييم أداء دالة التصنيف

بناء على بيانات العينة فان تقييم دالة التمييز سيكون كالتالي:

$$APER = \frac{n_{1m} + n_{2m}}{n_1 + n_2} = \frac{197 + 97}{567 + 233} = 0.367$$

حيث ان

n_{1m} = عدد العناصر من المجموعة الاولى المصنفة بشكل خاطئ في المجموعة الثاني ويساوي 197

n_{2m} = عدد العناصر من المجموعة الثانية المصنفة بشكل خاطئ في المجموعة الاولى ويساوي 97

اذن الخطأ الظاهر لتقدير دالة التصنيف كانت نسبته حوالي 37% الاستنتاجات

1. من خلال الدالة التمييزية الخطية تم التوصل إلى أكثر المتغيرات أهمية في التصنيف حيث كان متغير وزن الحقيبة إلى وزن الجسم أكثر أهمية ثم يليه متغير النوع ثم متغير العمر أكثر من باقي المتغيرات.
2. من خلال تصنيف البيانات وفق الدالة التمييزية الخطية وجد أن نسبة التصنيف الصحيح الكلي لها كان 63.3% في حين أن نسبة التصنيف الخاطئ كان 36.7%.
3. القدرة الجيدة للدالة التمييزية الخطية على التمييز بين المجموعتين. وبذلك فإن الاختلاف في التصنيف بين المجموعتين هو اختلاف معنوي ولا يعود للصدفة.

التوصيات

1. يجب على مسؤولي الدولة في قطاع التعليم الاهتمام بتلاميذ التعليم الأساسي والتركيز على الجانب الصحي والعوامل المؤثرة في هذه الجانب.
2. التأكيد على الوزن المثالي لحقيبة التلميذ وما يتناسب مع تكوينه البدني خاصة في المراحل الأولى من التعليم الأساسي.
3. إجراء دراسات مستمرة على تلاميذ المرحلة الأولى من التعليم الأساسي في كل المجالات خاصة منها الصحية والاقتصادية والاجتماعية والتي تؤثر في مسار العملية التعليمية للتلاميذ.

المراجع

1. أبو علام، رجاء (2003). التحليل الإحصائي للبيانات باستخدام برنامج SPSS. مصر، القاهرة: دار النشر للجامعات.
2. الحسن، أسماء ورحمة، عزيزة (2017). التحليل التمييزي لاستجابات عينة من الطلبة الجامعيين. مجلة جامعة البعث. المجلد 39، العدد 6 (103-144).
3. الجاعوني، فريد وعدنان غانم (2007). التحليل الإحصائي متعدد المتغيرات (التحليل التمييزي) في توصيف وتوزيع الأسر داخل الهيكل الاقتصادي الاجتماعي للمجتمع. مجلة جامعة دمشق للعلوم الاقتصادية والقانونية. المجلد الثالث والعشرون. العدد الثاني.
4. العباسي، عبد الحميد (2011)، تحليل التمايز. معهد الدراسات والبحوث الإحصائية، مصر.
5. جودة، محفوظ (2008)، التحليل الإحصائي الأساسي باستخدام SPSS. دار وائل للنشر، الطبعة الأولى، عمان، الأردن.
6. نجيب، حسين والرفاعي، غالب (2006)، تحليل ونمذجة البيانات باستخدام الحاسوب تطبيق شامللحزمة SPSS. الأردن، عمان: دار الأهلية للنشر والتوزيع.
7. يعقوب، أسماء (2017). التحليل العنقودي والتمييزي في دراسة تطبيقية على بعض المصارف العراقية. مجلة الاقتصاد الخليجي. العدد 31 (89-118).
8. Alvin C. Rencher (2002). "Methods of Multivariate Analysis" Second Edition, Brigham Young University.
9. Johnson R and Wichern D (2002). Applied Multivariate Statistical Analysis, Uppre Saddle River (NJ): Prentice-Hall

جدول رقم (6)

الدوال التمييزية الخطية التقديرية

| المتغير | الدالة الأولى | الدالة الثانية |
|---------------------------|---------------|----------------|
| النوع | 6.690 | 6.578 |
| العمر | 1.511 | 1.399 |
| الوزن | 0.252 | 0.240 |
| الطول | 0.034 | 0.035 |
| وزن الحقيبة | -0.374 | -0.580 |
| وزن الحقيبة إلى وزن الجسم | 8.521 | 7.880 |
| الثابت | -23.141 | -20.213 |

يوضح الجدول أعلاه رقم (6) معاملات الدوال التمييزية المعيارية والتي تحدد الأهمية النسبية لكل متغير من المتغيرات المستقلة التي تم ادخالها في عملة التحليل التمييزي. حيث نجد أن المتغير وزن الحقيبة إلى وزن الجسم هو أكثر أهمية في دالة التمييز يليه متغير النوع ثم متغير العمر، ثم متغير وزن الحقيبة يليه متغير الوزن وأخيرا متغير الطول في تشخيص متغير الشعور بالألم من عدمه. حيث يمكن كتابة الدوال التمييزية الخطية كما يلي:

$$Y_1 = 6.69X_1 + 1.511X_2 + 0.252X_3 + 0.034X_4 - 0.374X_5 + 8.521X_6 - 23.141$$

الدالة الثانية:

$$Y_1 = 6.758X_1 + 1.399X_2 + 0.240X_3 + 0.035X_4 - 0.580X_5 + 7.880X_6 - 20.213$$

تحديد النقطة الفاصلة والتحقق من التصنيف

بعد إيجاد الدالة التمييزية، يتم التحقق من تصنيف المشاهدات فيما إذا كانت فعلا تنتمي إلى المجموعة التي صنف لها. أي هل أن التصنيف كان صحيحا أم خاطئا. ولغرض إجراء ذلك لابد من تحديد النقطة الفاصلة بين المجموعات، حيث سيكون هناك نقطة فاصلة لكل مجموعة وكما هو موضح بالجدول رقم (7).

جدول رقم (7): النقاط الفاصلة لكل مجموعة

| الدالة | النقطة الفاصلة |
|--------|----------------|
| 1 | 0.187 |
| 2 | -0.455 |

والجدول رقم (8) يوضح دقة تصنيف المشاهدات

جدول رقم (8): نتائج دقة التصنيف

| | | الشعور بالألم | Predicted Group Membership | | Total |
|----------|-------|-------------------|----------------------------|---------------|-------|
| | | | عدم الشعور بالألم | الشعور بالألم | |
| Original | Count | عدم الشعور بالألم | 370 | 197 | 567 |
| | | الشعور بالألم | 97 | 136 | 233 |
| | % | عدم الشعور بالألم | 65.3 | 34.7 | 100 |
| | | الشعور بالألم | 41.6 | 58.4 | 100 |

63.3% of original grouped cases correctly classified من خلال جدول رقم (8) يتبين أن درجة صلاحية نموذج تصنيف المشاهدات (نسبة التصنيف الصحيح) بلغت 63.3%. كذلك نجد أن نسبة 65.3% من المشاهدات التي تنتمي إلى مجموعة عدم الشعور بالألم صنف بشكل صحيح إلى المجموعة التي تنتمي إليها، بينما نلاحظ أن نسبة 34.7% من المشاهدات صنف خطأ إلى مجموعة الشعور بالألم. في حين نجد أن نسبة 58.4% من المشاهدات التي تنتمي إلى مجموعة الشعور بالألم صنف بشكل صحيح إلى المجموعة التي تنتمي إليها، بينما نسبة 41.6% من المشاهدات صنف خطأ إلى مجموعة عدم الشعور بالألم.