مجلة العلوم التربوية والنفسية
JOURANL OF EDUCATIONAL AND PSYCHOLOGICAL SCIENCES

# استخدام نموذجي الانحدار اللوجستي والتحليل التمييزي لتحديد أهم العوامل المنبئة بالتعثر الدراسي لدى طلبة جامعة القصيم – دراسة مقارنة

**د. خالد محمد الحسن الحياصات**

أستاذ القياس والتقويم المساعد

**قسم علم النفس، كلية التربية، جامعة القصيم**

البريد الإلكتروني للباحث

km.alheasat@qu.edu.sa

تاريخ استلام البحث: ١ / ١٢ / ٢٠٢٣م

تاريخ قبول النشر: ٣١ / ١٢ / ٢٠٢٣م

# استخدام نموذجي الانحدار اللوجستي والتحليل التمييزي لتحديد أهم العوامل المنبئة بالتعثر الدراسي لدى طلبة جامعة القصيم – دراسة مقارنة

د. خالد محمد الحسن الحياصات

أستاذ القياس والتقويم المساعد

قسم علم النفس، كلية التربية، جامعة القصيم

**المستخلص:** هدفت الدراسة إلى المفاضلة بين أسلوبي تحليل الانحدار اللوجستي والتحليل التمييزي للكشف عن أبرز العوامل التي تُساهم في تصنيف الطلبة المتعثرين دراسياً، وتم استخدام محكات المقارنة بين الأسلوبين (دقة التصنيف، والمساحة تحت منحنى ROC)، ولتحقيق هذا الهدف تم استخدام المنهج الوصفي التحليلي، وتم الحصول على البيانات لعينة مكونة من (1068) طالباً وطالبة. حيث تم استخدام بيانات الطلبة كمتنبئات وهي (معدل الثانوية العامة، درجات الاختبار التحصيلي، درجات اختبار القدرات العامة، معدل التخصيص (معدل السنة التحضيرية)، أشارت نتائج الدراسة إلى أن ترتيب المتغيرات المستقلة حسب أهميتها النسبية في نموذج الانحدار اللوجستي هي (معدل الثانوية، درجات الاختبار التحصيلي، معدل التخصيص)، في حين كان ترتيبها وفق نموذج التحليل التمييزي (معدل الثانوية، معدل التخصيص، درجات الاختبار التحصيلي)، كما بينت نتائج الدراسة إلى أفضلية نموذج الانحدار اللوجستي في دقة التصنيف مقارنة بنموذج التحليل التمييزي، حيث بلغت النسبة المئوية لدقة التصنيف الصحيحة لنموذج الانحدار اللوجستي (97.8%)، في حين بلغت لنموذج التحليل التمييزي (90%)، إضافة إلى أن المساحة تحت منحنى ROC بلغت (99%) لأسلوب تحليل الانحدار اللوجستي، في حين بلغت لأسلوب التحليل التمييزي (97.7%).

**الكلمات المفتاحية:** الانحدار اللوجستي، التحليل التمييزي، التعثر الدراسي، معايير القبول

# Using logistic Regression and Discriminant Analysis Models for Detecting the Most Important Predictor Academic Failure Among Qassim University Students – A Comparative Study

**Dr. Khaled Mohammad Alhasan Alhyasat**

Assistant Professor of Measurement and Evaluation

*Department of Psychology - College of Education - Qassim University*

**Abstract**: The study aimed to compare the logistic regression analysis and the discriminatory analysis methods to Detecting the most important factors that contribute to the classification of students who are Failing academically. Criteria for comparison between the two methods (classification accuracy and area under the ROC curve) were used. To achieve this goal, the descriptive analytical method was used, and data was obtained for a sample of (1068) male and female students. Where the students' data was used as predictors, which are (General Secondary average, Achievement test scores, General Ability test scores, assignment rate (preparatory year average), The results of the study indicated that the order of the independent variables according to their relative importance in the logistic regression model is (General Secondary average, Achievement test scores, assignment rate), while their order according to the discriminant analysis model was (General Secondary average, assignment rate, Achievement test scores), as it showed The results of the study indicated the superiority of the logistic regression model in classification accuracy compared to the discriminant analysis model, as the percentage of correct classification accuracy for the logistic regression model was (97.8%), while it was (90%) for the discriminant analysis model, in addition to the area under the ROC curve reaching ( 99% for the logistic regression analysis method, while it reached 97.7% for the discriminant analysis method.

**Keywords**: logistic regression, discriminant analysis, academic failure, admission criteria.

**المقدمة**

يحظى التعليم الجامعي بأهمية كبيرة مرده عوامل ثقافية وعوامل اقتصادية وعوامل سياسية، ويعود ذلك إلى الأعداد الكبيرة من الطلبة الذين يلتحقون بالجامعات بهدف تحسين وضعهم العلمي والاجتماعي والمعيشي، ذلك أن التعليم يساهم بشكل كبير في تحسين كل ذلك (الأستاذ وصبح، ٢٠١٠).

وتشهد المملكة العربية السعودية تطوراً كبيراً في أعداد خريجي المرحلة الثانوية، قابل ذلك تزايد في الطلب على التعليم الجامعي، حيث يلتحق عشرات الآلاف من الطلبة بالجامعات سنوياً، ولعل ما تواجهه جامعات المملكة ومؤسسات التعليم العالي من تضخم في أعداد المتقدمين في كل عام، جعلها في تحد مستمر أمام اتخاذ قرار بقبول الطلبة أو رفضهم، لذلك كان لزاماً على المسؤولين استخدام معايير موضوعية تساعد إلى حد كبير في الوصول إلى قرار يتسم بقدر كبير من العدالة والصدق، ولذلك فقد استحدثت المملكة العربية السعودية معايير لقبول الطلبة كمعدل الثانوية العامة، واختبار القدرات العامة، والاختبار التحصيلي، ومعدل التحصيل (التخصيص) في السنة التحضيرية لتوزيع الطلبة داخل الكليات بالإضافة للمعايير الأخرى، وتعتمد إجراءات القبول في الجامعات السعودية على سياسة القبول التي تضعها كل جامعة وفقاً لشروط القبول الخاصة وتماشياً مع ما اعتمدته الدولة من معايير عامة (الخراشي، ٢٠٢١).

وتعد جامعة القصيم من أبرز الجامعات بالمملكة العربية السعودية التي يقصدها الكثير من خريجي الثانوية العامة لمتابعة تعليمهم الجامعي، وقد اعتمدت هذه المعايير ضمن شروطها لقبول الطلبة.

وعند تطبيق معايير القبول المتمثلة باختبار القدرات العامة والاختبار التحصيلي في الجامعات، لوحظ أن تسرب الطلبة وتعثرهم في التخصصات التنافسية قد انخفض بشكل لافت للنظر بعد تطبيق هذه الاختبارات؛ مقارنة بما كان عليه الوضع سابقاً، حيث كانت درجة الثانوية العامة وحدها هي المحك؛ لهذا لوحظ أن عدداً كبيراً من الطلاب قبل تطبيق اختبارات المركز الوطني للقياس والتقويم قد دخلوا تخصصات لا تتناسب وقدراتهم، وهذا أدى إلى فشلهم الذي امتد لسنوات، وانتهى بهم الحال بالانسحاب من الجامعة أو التخرج بعد الوقت المحدد بسنوات، وبمعدلات ضعيفة جداً لا تؤهلهم لسوق العمل (النعيمي، ٢٠٢٠).

وهناك العديد من المشكلات التي تؤدي إلى تعثر الطلبة أكاديمياً، وتتخذ هذه المشكلات صوراً متعددة، ولها العديد من الأسباب: فمنها ما يتعلق بشخصية الطالب، ومنها ما يتعلق بواقعه الأكاديمي، ومنها ما يتعلق بأسرته، ومنها ما يتعلق ببيئته، ومنها ما يتصل بحالة الطلاب الثقافية أو الاجتماعية، وتمثل هذه المشكلات نتيجة طبيعية لانشغال أولياء الأمور عن أبنائهم، ولأوجه القصور في المؤسسات التربوية النظامية وغير النظامية مثل البيت، والمدرسة، والجامعة، وغيرها من مؤسسات المجتمع الأخرى (صقر، ٢٠٠٣).

ويعتبر التعثر الدراسي من المشكلات الخطيرة التي تؤدي إلى عواقب وخيمة على الطالب والمجتمع، وأبرز نتائجها هو ترك مقاعد الدراسة والانقطاع عنها خاصةً في حالة تكرار الرسوب في الجامعة، وهذا ما قد يترك آثاراً سلبية على الطالب وأسرته وعلى المجتمع والتنمية بجميع أبعادها، بالإضافة إلى تأثيره نفسياً على الطالب، حيث يترك لديه مشاعر الخيبة والإحباط وعدم الرضا، مما يؤثر على تكيفه الاجتماعي، ودرجة انتمائه لمجتمعه، وقد يؤدي إلى تحويله إلى الانحراف والإجرام بدرجاتها المختلفة (العليان، ٢٠١٧).

وتعتبر مشكلة التنبؤ بنجاح الطالب من العوامل المهمة من المنظور العملي، حيث أن الأداء الأكاديمي للطالب لا يؤثر على قرارات قبوله من عدمها، وإنما يؤثر على اتجاهات الطلبة نحو التعليم والدراسة (Divjak & Oreski, 2009)، وتهتم كثير من الدراسات البحثية التطبيقية عند دراستها لظاهرة ما بتحديد المتغيرات المؤثرة في حدوثها والتنبؤ بها وخاصة عندما يكون المتغير التابع ثنائيًا، وتسعى في مثل هذه الحالات إلى استخدام أساليب تنبؤيه لها القدرة على التنبؤ بتصنيف الحالات وتمييزها (الشمراني، ٢٠١٧). وهناك العديد من هذه الأساليب الإحصائية التي تستخدم لأغراض التنبؤ وبصفة خاصة في مجال إعداد الاختبارات المنبئة التي يُعتمد عليها في إجراءات قبول الطلبة المتقدمين للالتحاق بمجال دراسي محدد أو وظيفة معينة، حيث أن هذه الاختبارات تستخدم لتحديد إمكانية نجاح الفرد أو فشله بالمجال الدراسي أو الوظيفة في حال التحاقه بأي منها (Wan & Hue, 2012). ويُعتمد في إعداد هذه الاختبارات كأدوات منبئة على قياس بعض المتغيرات التي لها علاقة بتلك المجالات الدراسية أو الوظيفية التي أثبتت الدراسات والتحليلات بأن لها القدرة في التنبؤ بحال الفرد بعد الالتحاق (William, 2010).

ويُعد كلاً من أسلوب الانحدار اللوجستي وأسلوب التحليل التميزي من الأساليب الإحصائية متعددة المتغيرات واسعة الاستخدام عند تحليل البيانات للمتغيرات التابعة التصنيفية، ويُطبق كل أسلوب منها على نطاق واسع في البحث وخاصة في المجال الطبي والعلوم الاجتماعية (Antonogeorgos et al., 2009)، وبالرغم من أن كل أسلوب يختلف عن الآخر في فكرته الأساسية إلا أن كليهما يعطيان نتائج لتصنيف المشاهدات (الشمراني، ٢٠١٧)، فالانحدار اللوجستي هو شكل من أشكال الانحدار الذي يستخدم عندما يكون المتغير التابع ثنائي التصنيف (Antonogeorgos et al., 2009)، وفي التحليل التمييزي يكون عدد الدوال التمييزية القانونية التي يتم تحديدها بعدد الفئات مطروح منها واحد، أو بعدد المتغيرات التمييزية الأصغر، فعلى سبيل المثال إذا كان هناك مجموعتان أو فئتان فإن الدالة التمييزية سيتم اشتقاقها لإعطاء أبسط شكل من أشكال التحليل التمييزي (Moawed & Osman, 2017).

وهناك اهتمام متزايد من قبل الباحثين لاختيار الأسلوب الإحصائي المناسب التي تتطابق فيها البيانات مع النموذج، حيث أنه من بين أكثر المعايير الأساسية للتمييز بين الطرق الإحصائية هو نوع متغير الاستجابة، وكذلك

الهدف من تصميم البحث، فإذا كان هناك متغير تابع تصنيفي ثنائي فإن النموذجين: الانحدار اللوجستي الثنائي والتحليل التمييزي الخطي يتم اقتراحهما كأحد النماذج متعدد المتغيرات الذين يستخدمان لتصنيف الحالات إلى مجموعاتها الأصلية، ويمكن استخدام هذه التقنية متعددة المتغيرات لمعرفة أي من المتغيرات المستقلة أفضل في التمييز بين مجموعتين أو أكثر لتصنيف الحالات إلى مجموعتها الطبيعية (Hamid, 2010).

وبالرغم من أن النظرية التي تخص كل من أسلوب الانحدار اللوجستي وأسلوب التحليل التمييزي منتشرة على نطاق واسع، لكن المقارنة بين هاتين الطريقتين لا زالت تُعد مشكلة للباحثين الذين يتقدمون للتمييز بين أثنين أو أكثر من النتائج التصنيفية، وبالتالي يمكن اقتراح أن كلاً من الانحدار اللوجستي والتحليل التمييزي الخطي يمكن استخدامهما في التنبؤ باحتمالية نتيجة محددة باستخدام كل المتغيرات المتاحة أو مجموعة فرعية من المتغيرات (Hamid, 2010; Timen, 2002)، ويمكن استخدام أي منهما وفق شروط يجب توافرها في المتغيرات التي يعتمد عليها في التنبؤ (Henrik, 2010)، فهناك العديد من الباحثين الذين يفضلون استخدام تحليل الانحدار اللوجستي؛ وذلك لأنه أقل شروط ولا يقدم أي افتراضات حول توزيع البيانات (المتغيرات المستقلة)، حيث يستخدم عندما لا تتطلب المتغيرات المستقلة افتراض التوزيع الطبيعي (Sharma, 1996)، كما أنه يعتبر أكثر قوة ومرونة في حال انتهاك هذه الافتراضات، بينما أسلوب التحليل التمييزي يتطلب تحقق افتراضات أساسية مثل التوزيع الطبيعي للبيانات، ولذلك من الضروري وضع توجيهات وإرشادات للاختيار الصحيح بين كلا الأسلوبين لأن هذا قد يؤدي هذا إلى نتائج قد تختلف، مما قد يؤثر ذلك سلباً على القدرة التنبؤية الخاصة بالاختبار (الشافعي، ٢٠١٣).

وعلى الرغم من وجود العديد من الدراسات السابقة التي اهتمت بالمقارنة بين الأسلوبين، إلا أنه مازال الخلاف قائماً بين أفضلية كل منهما، ولهذا الغرض جاءت هذه الدراسة لإجراء مقارنة بين أسلوبي التحليل التمييزي الخطي والانحدار اللوجستي من حيث أفضلية كل منهما في القدرة التنبؤية في تصنيف المشاهدات ثنائية التصنيف إذا ما تم استخدامها مع نفس المتغيرات وتحت نفس الظروف. وكذلك لتكشف عن أفضل العوامل التي تساهم في تصنيف الطلبة المتعثرين أكاديمياً وذلك باستخدام هذين الأسلوبين.

**مشكلة الدراسة**

يستخدم كلاً من أسلوب تحليل الانحدار اللوجستي وأسلوب التحليل التمييزي في تصنيف المشاهدات على مستوى متغير تصنيفي في ضوء بعض المتغيرات المستقلة، ولا يستلزم أسلوب تحليل الانحدار اللوجستي استخدام نفس شروط التحليل التمييزي فيما يتعلق بطبيعة توزيع المتغيرات المستقلة ومستوى قياسها، وحيث أن الاختلاف في تلك الشروط وعددها بين الأسلوبين قد تعطي قدرة تنبؤيه لأحدهما تختلف عن الآخر، وقد تكون القدرة التنبؤية لأحدهما في تصنيف المشاهدات أدق وأفضل من القدرة التنبؤية للآخر (الشافعي، ٢٠١٣)، بالتالي كانت هناك

حاجة ملحة لاستقصاء دقة أي من هذين الأسلوبين في عمليات التصنيف في ضوء بعض المتغيرات المستقلة، حيث أكدت بعض الدراسات على دقة أسلوب الانحدار اللوجستي على حساب التحليل التمييزي في التصنيف التنبؤي للمشاهدات ثنائية التصنيف منها دراسة حسينة (٢٠١٨)، ودراسة الرواشدة (٢٠٢٢)، ، ودراسة برزنجي (Barznji, 2018)، ودراسة إبراهيم زاده وزملاؤه (Ebrahimzadeh et al., 2015)، ودراسة بوهر وزملاؤه (Pohar et al, 2004)، ودراسة سانتوس وزملاؤه (Santos et al., 2014)، وعلى النقيض من ذلك فقد أثبتت بعض الدراسات العكس كدراسة كارديين وألموس (Kardiyen & Olmus, 2016)، ودراسة ليدي آني وجاستون (Lydi-Anne & Gaston, 2011)، ودراسة ماروكو وزملاؤه (Maroco et al., 2011)، في حين انتهت بعض الدراسات إلى وجود تشابه وتماثل بين أسلوبي التحليل موضع المقارنة من بينها: دراسة بورحاجي وآخرون (Poorhaji et al., 2022)، ودراسة موسى وآخرون (Musa et al., 2019)، ودراسة فيريرس وفيجو (Ferreras & Feijoo, 2009)، ودراسة إدلر وزملاؤه (Adler et al., 2011).

وقد لاحظ الباحث من خلال مراجعة الأدبيات السابقة التربوية والنفسية ندرة الدراسات العربية في هذا المجال، وأن جُل الأدبيات التي تناولت عقد المقارنة بين أسلوب تحليل الانحدار اللوجستي وأسلوب التحليل التمييزي كانت تتعلق بالمجالات الطبية والتجارية فقط دون المجالات التعليمية والتربوية وخاصة مجال تصنيف الطلاب المتعثرين دراسياً، ونظراً لأهمية الموضوع في المجالات التربوية بشكل عام، وتصنيف الطلبة في مختلف المجالات بشكل خاص ومنها تصنيف الطلبة المتعثرين دراسياً، كان لا بد من تقديم توصية بالأسلوب الأكثر دقة في تصنيف هؤلاء الطلبة، كما أن أغلب الدراسات السابقة التي حاولت الكشف عن الطلبة المتعثرين دراسياً استخدمت طرقاً وأساليب مختلفة للكشف على الطلبة المتعثرين منها: الاختبارات، ومعايير ديموغرافية، ومقاييس الخصائص السلوكية، فيما اقتصرت بعض الدراسات على استخدام أحد الأسلوبين كدراسة النعيمي (٢٠٢٠) التي اقتصرت دراسته على استخدام التحليل التمييزي، ولهذا الغرض جاءت هذه الدراسة لإجراء مقارنة بين أسلوبي التحليل التمييزي الخطي والانحدار اللوجستي من حيث أفضلية كل منهما في القدرة التنبؤية في تصنيف المشاهدات ثنائية التصنيف إذا ما تم استخدامها مع نفس المتغيرات وتحت نفس الظروف. وكذلك لتكشف عن أبرز العوامل التي تساهم في تصنيف الطلبة المتعثرين أكاديمياً وذلك باستخدام هذين الأسلوبين.

وبشكل أكثر تحديداً تحاول هذا الدراسة الإجابة عن التساؤلات الآتية:

**أسئلة الدراسة**

**السؤال الأول:** ما أبرز العوامل التي تساهم في التنبؤ بتعثر الطلبة دراسيًا باستخدام كل من أسلوب التحليل التمييزي، وأسلوب تحليل الانحدار اللوجستي؟

**السؤال الثاني:** ما مدى دقة كل من أسلوبي التحليل التمييزي وتحليل الانحدار اللوجستي في تصنيف الطلبة المتعثرين وغير المتعثرين دراسيًا؟

**السؤال الثالث:** أي الأسلوبين المستخدمين: التحليل التمييزي، الانحدار اللوجستي هو الأكثر دقة وفاعلية في تصنيف الطلبة المتعثرين دراسيًا؟

**أهداف الدراسة**

تهدف هذه الدراسة إلى إجراء مقارنة بين أسلوبين من أساليب التصنيف والتنبؤ، وهما التحليل التمييزي وتحليل الانحدار اللوجستي، من حيث ترتيب العوامل التي تساهم في تصنيف الطلبة المتعثرين دراسياً، وكذلك دقة التصنيف، وأي الأسلوبين المستخدمين أكثر فاعلية ودقة في التصنيف، وذلك لفهم كيفية عمل كلا الأسلوبين في الكشف عن العوامل التي تساهم بشكل معنوي في تصنيف الطلبة المتعثرين دراسياً.

**أهمية الدراسة**

تتمثل أهمية هذه الدراسة في جانبين أساسيين:

- **الجانب النظري:** يتمثل في تقديم معلومات جديدة للباحثين عن الأسلوب الأكثر دقة وفاعلية المستخدم في تصنيف الطلبة المتعثرين وغير المتعثرين دراسياً هل هو أسلوب الانحدار اللوجستي أم التحليل التمييزي؟، كذلك تقديم أهم العوامل التي تساهم في تصنيف الطلبة باستخدام كلا الأسلوبين.
- **الجانب التطبيقي:** يتمثل في مساعدة الباحثين في تحديد الأسلوب الأفضل لتصنيف الطلبة المتعثرين دراسياً، كما يكشف هذا البحث قدرة معايير القبول المستخدمة في جامعة القصيم في التنبؤ بالتعثر الدراسي، وتحديد المعايير الأكثر تمييزاً بين الطلبة المتعثرين وغير المتعثرين دراسياً.

**حدود الدراسة**

**الحدود الموضوعية:** اقتصرت الدراسة الحالية على دراسة أسلوبين في تحليل بيانات الطلبة وتصنيفهم، هما: أسلوب تحليل الانحدار اللوجستي الثنائي وأسلوب التحليل التمييزي، وعلى العوامل التي تتنبأ بالتعثر الدراسي للطالب في جامعة القصيم، وهي: معدل الثانوية العامة، درجات الاختبار التحصيلي، درجات اختبار القدرات، معدل التخصيص (المعدل التراكمي للسنة التحضيرية).

**الحدود المكانية:** جامعة القصيم في المملكة العربية السعودية.

**الحدود الزمانية:** الفصل الدراسي الأول للعام الجامعي 2022/2021م.

**الحدود البشرية:** الطلاب والطالبات الذين تجاوزوا السنة التحضيرية وتم تسكينهم في تخصصاتهم العلمية والانتهاء من دراسة السنة الأولى في التخصص.

**مصطلحات الدراسة**

**التحليل التمييزي:** يُعرّف بأنه: "أسلوب تحليل متعدد المتغيرات (Multivariate Analysis) يستخدم للتنبؤ بعضوية الحالة (أفراد، مشاهدات، ...) لمجموعات محددة ومعرّفة مسبقاً تمثل مستويات التنبؤ للمتغير التابع، وذلك من خلال تركيب خطي من متغير أو أكثر من المتغيرات المستقلة (المتنبئة)، وكذلك التنبؤ بالموقع التصنيفي لحالات جديدة لم تصنف بعد" (Huberty, & Olejnik, 2006, p.5).

**ويعرَّف في هذه الدراسة:** أسلوب احصائي يهدف إلى بناء نموذج تنبؤي يساعد على تصنيف الطلبة المتعثرين وغير المتعثرين بناء على درجاتهم في الثانوية العامة، والاختبار التحصيلي، واختبار القدرات العامة، ومعدله التراكمي في السنة التحضيرية (معدل التخصيص).

**تحليل الانحدار اللوجستي:** يُعرّف بأنه: "أسلوب إحصائي متعدد المتغيرات، يتم استخدامه لنمذجة العلاقة بين متغير تابع من المستوى الاسمي ومتغير مستقل أو أكثر من أي من مستويات القياس (الاسمي، الرتبي، الفئوي، النسبي)، بهدف تحديد احتمال عضوية مشاهدة لفئة معينة أو حصول حدث معين" (Hosmer et al., 2013, p.15).

**ويعرَّف في هذه الدراسة:** أسلوب احصائي يهدف للتنبؤ بنجاح الطالب أو تعثره بعد إنهائه للسنة التحضيرية وتسكينه بمجال دراسي معين اعتماداً على درجاته في الثانوية العامة، والاختبار التحصيلي، واختبار القدرات العامة، ومعدله التراكمي في السنة التحضيرية (معدل التخصيص).

**معايير القبول:** تُعرّف بأنها: "مجموعة من المقاييس التي تحددها الجامعة، وتشكل في مجموعها درجة كلية تكون عاملاً رئيسياً في تنافس الطلبة على المقاعد المتاحة" (الزامل، ٢٠١٢، ص.١٦١).

**وتعرَّف في هذه الدراسة:** هي المقاييس التي حددتها جامعة القصيم للمفاضلة بين المتقدمين للجامعة وهي: الاختبار التحصيلي، واختبار القدرات العامة، ومعدل الثانوية العامة.

**التعثر الدراسي:** يُعرّف بأنه: "عدم قدرة الطالب على إنجاز الساعات المقررة عليه خلال الفصول الدراسية المحدد لها من جهة أو تدني مستوى تحصيل الطالب كما يعكسه المعدل التراكمي له عن المعدل المقبول (65%) من جهة أخرى، الأمر الذي قد يترتب عليه حصول الطالب على إنذار أكاديمي أو وقف لقيده أو إجباره على التحويل من كلية إلى كلية أخرى من جهة ثالثة" (الأستاذ وصبح، ٢٠١٠، ص.٤٤).

**ويعرَّف في هذه الدراسة:** بحصول الطالب على معدل تراكمي أقل من (٢) من (٥) في الفصل الدراسي الأول من السنة التي تلي السنة التحضيرية من دراسته في جامعة القصيم.

## الإطار النظري والدراسات السابقة

### الإطار النظري

يعد أسلوبي التحليل التمييزي والانحدار اللوجستي من الأساليب الإحصائية متعددة المتغيرات واسعة الاستخدام عند تحليل البيانات للمتغيرات التابعة التصنيفية من أجل تصنيف الحالات أو المتغيرات، فكلاهما مناسب لنماذج التصنيف الخطية، ويختلف الأسلوبان في فكرتهما الأساسية، فأسلوب الانحدار اللوجستي لا يتطلب افتراضات تتعلق بتوزيع البيانات، بينما يتطلب أسلوب التحليل التمييزي تحقق افتراضات أساسية مثل افتراض التوزيع الاعتدالي، ولكنهما يتوصلان إلى نتائج لتصنيف المشاهدات (Pohar et al., 2004).

### نموذج الانحدار اللوجستي Logistic Regression Model

يتم استخدام نموذج الانحدار اللوجستي الثنائي عندما يأخذ المتغير التابع (متغير الاستجابة) قيمتين فقط. يستخدم هذا النموذج بشكل أساسي لتحديد العلاقة بين واحد أو أكثر من المتغيرات المستقلة (Xi) والمتغير التابع (Y)، أو للتنبؤ بالمتغيرات المستقلة الأكثر تأثيرًا على المتغير التابع.

يحدد شكل نموذج الانحدار اللوجستي العلاقة بين احتمال الاستجابة ومتغيرات التوقع على النحو الآتي:

$$Logit(P_i) = \log\left(\frac{P_i}{1-P_i}\right) = \beta_{i0} + \beta_{i1}X_{i1} + \cdots + \beta_{ik}X_{ik}$$

حيث، $\left(\frac{P_i}{1-P_i}\right)$ هي نسبة احتمال النجاح إلى احتمال الفشل، وتسمى معامل الترجيح أو المفاضلة Odds، β0، βi هي المعلمات التي سيتم تقديرها، X المتغير المستقل، وpi هو احتمال الاستجابة للمجموعة i، k هو عدد المتغيرات (Alkarkhi et al., 2008).

### افتراضات نموذج الانحدار اللوجستي Assumption Logistic Regression Model

هناك عدة افتراضات لنموذج الانحدار اللوجستي وهي كالآتي (Tabachnich & Fidell, 2013):

١. المتغير التابع هو متغير نوعي تصنيفي يأخذ فئتين أو أكثر.
٢. المتغيرات التفسيرية التي يشملها النموذج متغيرات مستقلة عن بعضها البعض.
٣. العلاقة بين المتغير التابع والمتغيرات المستقلة تكون علاقة دالة غير خطية.
٤. حد الخطأ يتبع توزيع ذو حدين (Binomial Distribution) بتوقع يساوي 0 وبتباين يساوي 1.
٥. وجود حد أدنى من التداخل الخطي المتعدد (Multicolinearity) بين المتغيرات المستقلة أو عدم وجودها، وبالتالي فإنه يجب حذف المتغيرات التي بها علاقة تامة.
٦. حجم العينة، حيث يتطلب الانحدار اللوجستي عادة حجم عينة كبير.

**طرق الانحدار اللوجستي**

هناك ثلاثة طرائق لإدخال المتغيرات في الانحدار اللوجستي كالآتي (Tabachnich & Fidell, 2013):

١. **الانحدار اللوجستي المباشر Direct Logistic Regression**: يتم ادخال المتغيرات إلى التحليل دفعة واحدة دون إعطاء أي أهمية لأي متغير.

٢. **الانحدار اللوجستي التسلسلي Sequntial Logistic Regression** وفي هذا النوع يتم ادخال المتغيرات واحداً تلو الآخر.

٣. **الانحدار اللوجستي التدريجي Stepwise Logistic Regression** يتم ادخال المتغيرات للتحليل حسب معيار إحصائي يحدد أولوية إدخال المتغيرات إلى النموذج.

**التحليل التمييزي الخطي Linear Discriminant Analysis**

يتنبأ التحليل التمييزي الخطي بمتغير تابع فئوي باستخدام متغيرات مستقلة مستمرة أو ثنائية. الدالة التمييزية هي مجموعات خطية من المتغيرات المفيدة عند تحديد ما إذا كانت مجموعة من المتغيرات فعالة للتنبؤ بعضوية الفئة. من المفترض أن المتغيرات لها توزيع طبيعي متعدد المتغيرات وأن مصفوفات التباين/التغاير للمتغيرات متجانسة عبر المجموعات. شكل الدالة التمييزية هو:

$$Z_i = a_i + w_{i1}X_{i1} + w_{i2}X_{i2} + \cdots + w_{ik}X_{ik}$$

في هذه الصيغة، يشير الحرف i إلى المجموعة المحددة؛ تشير الأحرف 1، 2، ...، k إلى متغيرات k؛ $a_i$ ثابت، $w_{ij}$، j=1، ...، k هو وزن المتغير j في حساب درجة التصنيف للمجموعة i؛ Xj هي القيمة المرصودة للمتغير jth. Zi هي درجة التصنيف. تحسب هذه الصيغة درجات التصنيف لكل حالة. يتم بعد ذلك تصنيف الحالات إلى مجموعات محددة بناءً على أعلى درجات التصنيف.

**افتراضات التحليل التمييزي الخطي Assumption Linear Discriminant Analysis**

يستند التحليل التمييزي الخطي إلى عدد من الافتراضات التي تحدد إمكانية الاعتماد على دواله التمييزية للتصنيف بالدقة المطلوبة وهي (Gilbert, 1969):

١. التوزيع الطبيعي: يتطلب التحليل التمييزي الخطي أن تتوزع المتغيرات المتنبية توزيعا طبيعيا لكل من مستويات المتغير التصنيفي، ويتم التحقق من هذا الافتراض من خلال اختبار كولموغروف سميرنوف (kolmogorov-Smirmov).

٢. تجانس مصفوفة التغاير والتباين: أي أن تتساوي التباينات والتغايرات عبر مستويات المتغير التصنيفي، ويتم التحقق من هذا الافتراض من خلال اختبار بوكس (Box's-M).

٣. الاستقلالية: وتعني عدم تأثر اختبار أي عنصر من العينة بعنصر آخر، ولتحقيق الاستقلالية يتم اختبار العينة بشكل عشوائي.

**أنواع التحليل التمييزي:**

في حال تم النظر إلى طريقة استخدام المتغيرات المستقلة في هذا التحليل؛ فإنه يمكن القول بأن هناك ثلاثة أنواع هي الآتي (William, 2010):

١. **التحليل التمييزي المباشر Direct Discriminant Analysis**: وفي هذا النوع من التحليل يتم استخدام المتغيرات المستقلة دفعة واحدة عند إنشاء دالة أو دوال التمييز (يتم ادخال المتغيرات إلى التحليل دفعة واحدة دون إعطاء أي أهمية لأي متغير).

٢. **التحليل التمييزي الهرمي hierarchical Discriminant Analysis**: وفي هذا النوع يتم استخدام المتغيرات المستقلة – أثناء إنشاء دالة أو دوال التحليل التمييزي – بالترتيب وليس دفعة واحدة بحيث يتم استخدام أحد المتغيرات ثم يتم الانتقال إلى المتغير الثاني وهكذا، وذلك وفقاً للترتيب الذي اختاره الباحث أو المستخدم. (يتم ادخال المتغيرات حسب رؤية الباحث).

٣. **التحليل التمييزي التدريجي stepwise Discriminant Analysis**: هو شبيه بالنوع السابق إلا أن ترتيب المتغيرات فيه يكون وفقاً لمعايير إحصائية، حيث يتم البدء بالمتغير الأكثر تمييزاً بين المجموعات. يتم ادخال المتغيرات للتحليل حسب معيار إحصائي يحدد أولوية إدخال المتغيرات إلى النموذج حيث يتم إضافة المتغيرات إلى الدوال التميزية واحد تلو الآخر حتى نجد أن إضافة مُتغيرات لا يُعطي تمييزا أفضل).

٤. أما فيما يتعلق بعدد الدوال التمييزية التي ينتجها هذا التحليل فإن هناك نوعين من التحليل التمييزي، هما:

أ. **التحليل التمييزي الأحادي** (DA): وهو التحليل الذي يحتوي على دالة تمييزية واحدة.

ب. **التحليل التمييزي المتعدد** (MDA): وهو التحليل الذي يستخدم أكثر من دالة تمييزية.

**تقييم مطابقة النموذج**

يمكن التحقق من مطابقة بيانات النموذج المستخرجة من عينة التحليل باستخدام عدة طرق منها:

**أولاً: طريقة البواقي (Residual)**: من خلال هذه الطريقة يتم التأكد من وجود دلالة إحصائية للمتغيرات المستقلة في النموذج ومدى ارتباطها بالمتغير التابع باستخدام الصيغ (Kutner et al., 2005):

$$person\ residual = e_i = \frac{y_i - n_i\hat{\pi}_i}{\sqrt{n_i\hat{\pi}_i(1-\hat{\pi}_i)}}$$

$$Standardized\ residual = e_i = \frac{y_i - n_i\hat{\pi}_i}{SE}$$

حيث $SE = \sqrt{n_i\hat{\pi}_i(1-\hat{\pi}_i)(1-h_{ii})}$، $n_i$: عدد الحالات، $y_i$: قيم الخطأ المقدرة، $\hat{\pi}_i$: احتمال القيمة المتوقعة، $h_{ii}$: العناصر القطرية للمصفوفة $n \times n$، $i: 1.2.\cdots.n$

**ثانياً: طريقة جدول التصنيف Classification Table**

من هذه خلال الطريقة يتم مقارنة النتائج المشاهدة مع النتائج المتوقعة باستخدام النموذج، من خلال قيمة احتمالية متوقعة (cut value)، وهي غالباً تساوي (0.50) (Josef, 2009). ومن هذه الطريقة يمكن تقدير الحساسية والنوعية والدقة (نسبة التصنيف الصحيح) كالآتي:

❖ **الحساسية Sensitivity**: وهي احتمال أن يكون التصنيف المتوقع موجباً للمشاهدات الموجبة؛ أي قدرة التحليل على تحديد الطلبة المتعثرين بشكل صحيح أنهم فعلاً متعثرين، ويمكن حسابها من خلال الصيغة الآتية:

$$\text{Sensitivity} = \frac{TP}{TP + FN} = \frac{TP}{P}$$

حيث: P المشاهدات الموجبة في التصنيف، TP المشاهدات المتوقعة الصحيحة، FN المشاهدات المتوقعة السالبة الخاطئة.

❖ **النوعية/القابلية Specificity**: وهي احتمال أن يكون التصنيف المتوقع سالباً للمشاهدات التي تكون فعلاً سالبة؛ أي قدرة أو قابلية التحليل على تحديد الطلبة غير المتعثرين بشكل صحيح أنهم فعلاً غير متعثرين. وتحسب من خلال الصيغة الآتية:

$$\text{Specificity} = \frac{TN}{FP + TN} = \frac{TP}{P'}$$

حيث: $TN$ المشاهدات المتوقعة السالبة الصحيحة، $P'$: المشاهدات السالبة في التصنيف.

❖ **الدقة Accuracy** أو **نسبة التصنيف الصحيحة Hit Ratio**: وهي عبارة عن قيمة احتمال التصنيف الصحيح، فإذا كانت الكفاءة (Efficiency) والتي تأخذ الصيغة $Ef = TP + TN$ فإن نسبة التصنيف الصحيح تساوي:

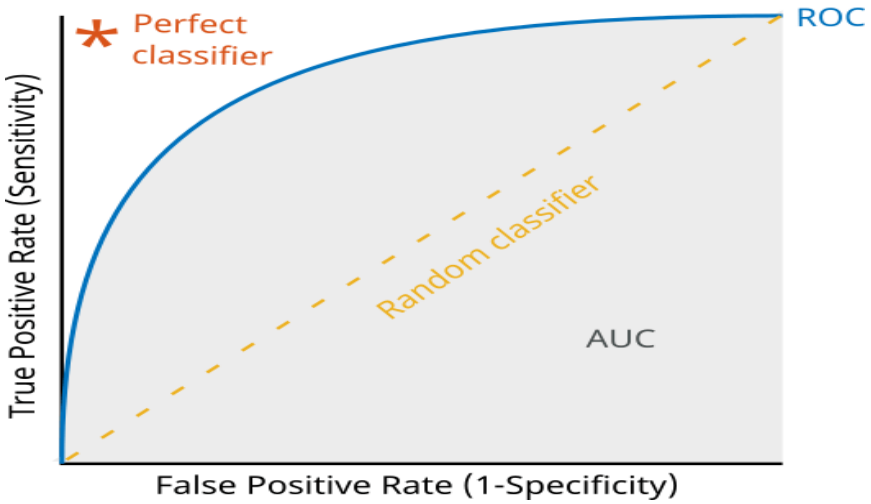
$$\text{Hit Ratio} = \frac{Ef}{Total} = \frac{TP + TN}{TP + FP + TN + FN} = \frac{TP + TN}{P + P'}$$

**ثالثاً: طريقة المساحة تحت منحنى ROC: AUC (Area Under Curve)**

يمثل منحنى ROC (**Receiver Operating Characteristic curve**) رسماً بيانياً يوضح أداء نموذج التصنيف، حيث يمثل الحساسية **Sensitivity** مقابل (1 – النوعية) (**1–Specificity**)، كم يوضحه الشكل رقم (١):

**الشكل ١**

***المساحة تحت منحنى ROC***




ويتم حساب المساحة تحت منحنى ROC على النحو الآتي:

$$AUC = \sum_{i=1} S_i$$

$$S_i = TFP_i - TFP_{i-1} \times \frac{(TVP_i - TVP_{i+1})}{2}$$

حيث تشير $TVP$ إلى نسبة الموجب الصحيح، $TFP$ نسبة الموجب الخاطئة

وأشار هوزمير وآخرون (Hosmer et al., 2013) إلى أن تفسير قيمة المساحة تحت المنحنى يكون كالآتي: إذا كانت $AUC = 0.50$ يكون النموذج ليس له قدرة تصنيفية تختلف عن الصدفة (أي النموذج غير قادر على تصنيف الطلبة المتعثرين والغير متعثرين)، $0.60 \leq AUC < 0.70$، تكون قدرة تفسيرية غير دقيقة (سيئة)، $0.60 \leq AUC < 0.70$ تكون قدرة تفسيرية ضعيفة، بينما إذا كانت $0.70 \leq AUC < 0.80$ تكون قدرة تفسيرية مقبولة، وإذا كانت $0.80 \leq AUC < 0.90$ تكون قدرة تفسيرية جيدة، وإذا كانت $0.90 \leq AUC$ تكون قدرة تفسيرية ممتازة.

**رابعاً: طريقة اختبار الدرجة Score Test**: هذا الاختبار لا يتطلب الاعتماد على نسبة الأرجحية العظمى (Hosmer et al., 2013)، وإنما باستخدام الدرجات الخام والذي يتم حسابه بالصيغة الآتية:

$$ST = \frac{\sum_{i=1}^{n} x_i(y_i - \bar{y})}{\sqrt{\bar{y} - (1 - \bar{y}) \sum_{i=1}^{n} (x_i - \bar{x})}}$$

حيث $x_i$: ترميز لقيم لمشاهدات في المتغير المستقل، و $y_i$: ترميز لقيم لمشاهدات في المتغير التابع.

**خامساً: طريقة إحصاء والد (Wald Statistics)**: يستخدم هذا الاحصائي للتعرف هل كل متغير من المتغيرات المستقلة له تأثير معنوي على المتغير التابع في النموذج أم لا؟ ويستخدم هذا الاختبار لاختبار الدلالة الإحصائية لمعاملات انحدار نموذج الانحدار اللوجستي، والذي له توزيع $\chi^2$ بدرجة حرية واحدة. ويحسب بالصيغة الاتية:

$$Wald\ Statistics = \left(\frac{\hat{B}}{SE(\hat{B})}\right)^2$$

حيث $\hat{B}$: هي تقدير الأرجحية العظمى لمعامل الانحدار B ، $SE_b$: هي تقدير الخطأ لمعامل الانحدار B.

**الدراسات السابقة**

أجرى بورهاجي وآخرون (Poorhaji et al., 2022) دراسة هدفت إلى مقارنة دقة التحليل التمييزي الخطي والانحدار اللوجستي الثنائي في التنبؤ بمرض فايروس كورونا (كوفيد – 19) الشديد وغير الشديد، واستخدمت الدراسة بيانات حقيقية تكونت العينة من (١٠٩) مرضى مصابين بالتهاب رئوي مؤكد بسبب كوفيد – 19، تكونت المجموعة المصابة إصابة خفيفة غير حادة من (42) مريضاً، والمجموعة الثانية المصابة إصابة شديدة من (٦٧) مريضاً. أظهرت نتائج الدراسة فاعلية الأسلوبين والتشابه بينهما في تصنيف المرضى الذين يعانون من حالات كوفيد-19 الحادة وغير الشديدة بشكل فعال.

وأجرت الرواشدة (٢٠٢٢) دراسة هدفت إلى المفاضلة بين أسلوبي التحليل التمييزي والانحدار اللوجستي لدراسة أهم العوامل التي تساهم في تصنيف الطلبة إلى موهوبين ومتفوقين، وتكونت عينة الدراسة من (٥١٤) طالباً وطالبة في مدارس إربد، وبينت نتائج الدراسة أن النسبة المئوية لدقة التصنيف لأسلوب تحليل الانحدار اللوجستي بلغت (95.9%) مقابل (92.6%) لأسلوب التحليل التمييزي، كما بينت النتائج أن المساحة تحت المنحنى (ROC) لأسلوب الانحدار اللوجستي بلغت (98.7%)، بينما بلغت لأسلوب التحليل التمييزي (97.5%)، وبيّنت نتائج الدراسة أفضلية أسلوب تحليل الانحدار في دقة التصنيف مقارنة بالتحليل التمييزي في تصنيف الطلبة الموهوبين.

وفي دراسة محاكاة قام بها البنوي (٢٠٢١) للمقارنة بين فاعلية كل من أسلوبي التحليل التمييزي والانحدار اللوجستي بدلالة دقة التصنيف في ظل الظروف الاختبارية (حجم العينة، عدد التنبئات، عدد مستويات التنبؤ، حجم الأثر)، كما هدفت إلى التنبؤ بالطريقة الأكثر فاعلية في ظل الظروف الاختبارية، وتم استخدام بيانات تم توليدها لمختلف الظروف الاختبارية. بيّنت نتائج الدراسة أن الانحدار اللوجستي أكثر فاعلية من التحليل التمييزي في ظل كل من: حجم العينات الأصغر، وعدد المتنبئات الأكبر، وكذلك عندما يكون التصنيف في مجمعتين إلى أربع مجموعات، بينما كان التحليل التمييزي أكثر فاعلية من الانحدار اللوجستي عند التصنيف في خمس مجموعات وأكثر.

وقام موسى وآخرون (Musa et al., 2019) بدراسة هدفت إلى تقييم التقارب بين أسلوب التحليل التمييزي الخطي والانحدار اللوجستي عندما يتم تطبيقها في بيانات وبائية غير موزعة بشكل طبيعي، وقد استخدمت الدراسة عدة معايير لمقارنة الأسلوبين هي: معاملات كل نموذج، وتأثير حجم العينة على النسبة المئوية لتصحيح التصنيف، والحساسية والنوعية، ودقة النماذج والمنطقة الواقعة تحت منحنى (ROC) باستخدام تحليل منحنى (ROC) الذي أجري بشكل فردي باستخدام الاحتمالات المتوقعة المحفوظة من الأسلوبين الإحصائيين، وبينت نتائج الدراسة أن كلا أسلوبي التحليل نموذج الانحدار اللوجستي والتحليل التمييزي أعطت نتائج متماثلة في التصنيف حتى مع انتهاك افتراض التوزيع الطبيعي.

وقامت برزنجي (Barznji. 2018) بدراسة هدفت إلى المقارنة بين تحليل الانحدار اللوجستي والتحليل التمييزي الخطي لاختيار أفضل نموذج إحصائي لتحديد عوامل الخطر المرض السكري. واشتملت الدراسة على متغير تابع تصنيفي (وجود مرض، عدم وجود مرض) وسبعة متغيرات مستقلة (المتنبئات) وقد تم مناقشة نتائج كل من تحليل الانحدار اللوجستي والتحليل التمييزي الخطي بعد تطبيقهما على بيانات عينة تكونت من (250) مريضاً من مرضى السكري، وقد أظهرت نتائج هذه الدراسة أن تحليل الانحدار اللوجستي لديه أفضل تنبؤ لعوامل الخطر لمرض السكري.

كما أجرت حسينة (٢٠١٨) دراسة هدفت إلى المقارنة بين نموذج الانحدار اللوجستي وأسلوب التحليل التمييزي لتقييم تنبؤ هذه الطرق بفشل المؤسسات الاقتصادية، وتم استخدام معيار التقييم الإحصائي ونسبة التصنيف الصحيح الكلي ومساحة تحت المنحنى (ROC) ، وبينت النتائج أن نموذج الانحدار اللوجستي له أفضلية في التنبؤ بفشل المؤسسات الاقتصادية من أسلوب التحليل التمييزي، كما كان تغير واضح في نسب التصنيف الكلي باستخدام معيار الصلاحية المقطعية، حيث أكدت نتائجه أن نموذج الانحدار اللوجستي والتحليل التمييزي لها نفس دقة التصنيف الكلي إلا أن الانحدار اللوجستي تفوق على التحليل التمييزي في التنبؤ بالمؤسسات الفاشلة.

وأجرى الشمراني (٢٠١٧) دراسة هدفت إلى المقارنة بين أسلوبي انحدار الترجيح اللوغاريتمي وتحليل الدالة التمييزية في تحديد مجموعة المنبئات التي تميز بدرجة أفضل بين الأداء الأكاديمي المرتفع والمنخفض ودقة تصنيف المجموعتين. وتكونت عينة الدراسة من (١٧٢٢) طالباً وطالبة بجامعة أم القرى، واستخدم الباحث المتغيرات المستقلة الآتية: المعدل التراكمي للطلبة في السنة التحضيرية، ودرجات الاختبار التحصيلي، ودرجات اختبار القدرات، ودرجات الثانوية، كما استخدم متغير الأداء الأكاديمي المنخفض والمرتفع كمتغير تابع. وأشارت نتائج الدراسة إلى أن حساسية التصنيف باستخدام أسلوب تحليل انحدار الترجيح اللوغاريتمي بلغت (84.4%)، بينما بلغت لأسلوب تحليل الدالة التمييزية (83.7%) بفرق (0.7) لصالح تحليل انحدار الترجيح اللوغاريتمي، كما أشارت النتائج إلى أن دقة التصنيف عند استخدام تحليل انحدار الترجيح اللوغاريتمي بلغت (87.3%)، في حين بلغت (87%) عند

استخدام تحليل الدالة التمييزية وبفرق (0.3) لصالح تحليل انحدار الترجيح اللوغاريتمي، وأشارت النتائج أيضاً إلى أن قيمة نسبة التصنيف الصحيح عند استخدام أسلوب تحليل انحدار الترجيح اللوغاريتمي بلغت (85.9%)، بينما كانت قيمتها عند استخدام تحليل الدالة التمييزية (85.4%) وبفرق (0.4) لصالح تحليل انحدار الترجيح اللوغاريتمي. وأشارت النتائج إلى تشابه نتائج تحليل انحدار الترجيح اللوغاريتمي وتحليل الدالة التمييزية في تحديد العوامل التمييزية وقوتها التنبؤية والتي تُسهم في تصنيف الطلبة إلى مرتفعي ومنخفضي التحصيل مرتبة حسب أهميتها ومساهمتها: متغير معدل السنة التحضيرية، ثم متغير درجات الثانوية، ثم متغير الاختبار التحصيلي، والأقل متغير اختبار القدرات العامة.

كما أجرى كارديان وأُلموس (Kardiyen & Olmus, 2016) دراسة هدفت إلى المقارنة بين تصنيف مجموعتين باستخدام ثلاث طرق، هي: التحليل التمييزي والانحدار اللوجستي المتعدد وتحليل التباين المتعدد. وتم إجراء عملية المقارنة باستخدام عينات محاكاة لمختلف التوزيعات وأحجام العينات، وأظهرت نتائج الدراسة أن التحليل التمييزي كان ذو أفضلية في تصنيف البيانات عندم تقترب توزيعات المجموعات من التوزيع الطبيعي.

كما سعت دراسة ميهالوفيتش (Mihalovic, 2016) لتطوير نموذج للتنبؤ بإفلاس الشركات في الجمهورية السلوفاكية، وركزت الدراسة على المقارنة في أداء التنبؤ لكلا النموذجين التحليل التمييزي والانحدار اللوجستي، ولقد تكونت العينة من (٢٣٦) شركة تعمل في سلوفاكيا مقسمة إلى مجموعتين (الشركات الفاشلة والشركات غير الفاشلة)، وأظهرت النتائج أن أهم المتنبئات الدالة التي تعيق فشل الشركات هي (صافي الدخل إلى أجمالي الأصول، والنسبة الجارية، والخصوم المتداولة إلى إجمالي الأصول)، كما أظهرت نتائج الدراسة إلى أن نموذج الانحدار اللوجستي تفوق في دقة التصنيف على أسلوب التحليل التمييزي.

وأجرى كل من طاقية وآخرون (٢٠١٦) دراسة هدفت إلى استخدام أسلوبي الانحدار اللوجستي والتحليل التمييزي لدراسة حالات الإصابة بمرض الإسهال لدى الأطفال في العراق، وتم تصنيف البيانات وفق ثلاثة نماذج: الأول نموذج الدالة الاحتمالية لدالة الانحدار اللوجستي المتعددة الاستجابة، والثاني نموذج دالة التمييز الخطي، والثالث نموذج الاحتمالات الاستجابة لدالة التمييز الخطي، واستخدم الباحث أسلوبين من أساليب التصنيف والتنبؤ هما التحليل التمييزي والانحدار اللوجستي وعمل مقارنة بينهما لتحديد أهم العوامل الذي تؤثر على طبيعة الإصابة بمرض الإسهال لدى الأطفال في العراق، وقد تكون هذه العوامل اقتصادية وديموغرافية وطبية وغيرها والمتمثلة بالمتغيرات المستقلة وما مدى تأثيرها على المتغير التابع ومعرفة أي من النماذج الثلاثة أفضل وأدق في توظيف البيانات وأيهما يظهر بأقل نسبة خطأ تصنيف ممكن، وقد أظهرت نتائج الدراسة أن طريقة التصنيف بحسب الانحدار اللوجستي أعطى أقل احتمال خطا تصنيف، في حين أعطت الصيغة الاحتمالية لدالة التمييز الخطي أعطت نتائج أكثر احتمال

خطأ التصنيف، وبينت الدراسة إلى أن تصنيف البيانات بطريقة الصيغة الاحتمالية لنموذج الانحدار اللوجستي تفوقت على دالة التمييز الخطي وذلك بإعطائها أقل احتمال لخطأ التصنيف.

كما أجرى شايان وآخرون (Shayan et al., 2015) دراسة هدفت إلى التنبؤ بالاكتئاب لدى مرضى السرطان باستخدام معايير تصنيف مختلفة، وتكونت عينة الدراسة من (243) من مرضى السرطان في أقسام العلاج الكيميائي والعلاج الإشعاعي في إحدى المستشفيات بأحجام عينات مختلفة، وتم حساب مؤشرات التصنيف (CE, Q, B) باستخدام نموذج الانحدار اللوجستي والتحليل التمييزي. أظهرت نتائج الدراسة عدم تفوق أحد النماذج على الآخر عند استخدام مؤشر خطأ التصنيف (CE)، بينما كان أداء نموذج الانحدار اللوجستي أفضل من التحليل التمييزي في جميع المواقف عند استخدام مؤشري التصنيف Q و B. وبالتالي يعد نموذج الانحدار اللوجستي LR مناسبًا للتنبؤ بالاكتئاب لدى مرضى السرطان بناءً على متغير الرضا عن حالته، والتاريخ العائلي للاكتئاب، ومدة الإصابة بالسرطان باستخدام مؤشرات Q و B.

وأجرى إبراهيم زاده وآخرون (Ebrahimzadeh et al., 2015) دراسة هدفت إلى التنبؤ بحالات الحمل المرغوب وغير المرغوب فيه. وقد قارنت الدراسة بين استخدام نموذج الانحدار اللوجستي والانحدار الاحتمالي والتحليل التمييزي، واستخدمت الدراسة ثلاثة نماذج للتصنيف، تكونت العينة من (787) من النساء الحوامل في عام 2012، وتم استخدام عينة طبقية وعنقودية، واستخدمت محكات الحساسية، ونسبة التوقعات. وتوصلت نتائج الدراسة إلى أن دقة التصنيف باستخدام أسلوب نموذج الانحدار اللوجستي كانت أفضل من أسلوب التحليل التمييزي.

كما أجرى سانتوس وآخرون (Santos et al., 2014) دراسة تتعلق بتحديد نوع الجنين وفق قياسات الجمجمة باستخدام ثلاث طرق منها أسلوبي التحليل التمييزي الخطي ونموذج الانحدار اللوجستي، واستخدمت الدراسة عينات مختلفة من عدة دول للمقارنة بين هذه الطرق الإحصائية، وأجريت المقارنة باستخدام محكات مختلفة، وأشارت النتائج إلى أن دقة التنبؤ باستخدام نموذج الانحدار اللوجستي كانت أفضل من التحليل التمييزي.

وأجرى كل من ليدي آن وجاستون (Lydi-Anne & Gaston, 2011) دراسة للمقارنة بين أسلوب التحليل التمييزي وتحليل الانحدار اللوجستي الثنائي لبعض العوامل بالتنبؤ بعدم شرب الكحول أثناء فترة الحمل والتي تمثلت في معتقدات الحامل نحو مضار شرب الكحول على الطفل المولود والعمر للحامل أثناء فترة الحمل. من أجل تحقيق ذلك تم أخذ بيانات لعينة عشوائية من الحوامل بلغت (167) امرأة في سن الانجاب، وتم تحليل البيانات باستخدام أسلوب التحليل التمييزي وتحليل الانحدار اللوجستي الثنائي، وقد توصلت نتائج الدراسة إلى أن أسلوب التحليل التمييزي كان أكثر دقة في تصنيف متعاطي الكحول وغير المتعاطين مقارنة بأسلوب الانحدار اللوجستي الثنائي.

وفي دراسة أجراها جريجوري وآخرون (Gregori et al., 2011) بهدف تحديد العوامل التنبؤية بموت مرضى السكر، وقد استخدمت الدراسة عدة اساليب إحصائية تنبؤية منها التحليل التمييزي وتحليل الانحدار اللوجستي، وقد تم اخذ عينة من سجلات بيانات مرضى السكر بأحد المستشفيات في مدينة تورينو بايطاليا بلغت (3892) مريضاً والذين تم تشخيص إصابتهم بداء السكر النوع الثاني، وشملت هذه العوامل عمر المريض، تاريخه المرضي، نوع الغذاء، وطبيعة العمل. وقد بيَّنت نتائج الدراسة أن أسلوب الانحدار اللوجستي كان أكثر دقة في تحديد العوامل المنبئة بموت مرض السكر وذلك مقارنة بالتحليل التمييزي الخطي.

كما أجرى ماروكو وآخرون (Maroco et al., 2011) دراسة تنبؤية للتنبؤ بالخرف، واستخدمت عدة طرق تنبؤية ومنها التحليل التمييزي وتحليل الانحدار اللوجستي. وقد استخدمت الدراسة بيانات حقيقية لعينة مرضى تكونت من (٩٢) حالة من أجل تصنيف هذه الحالات إلى ضعف إدراكي معرفي وغير ذلك، حيث اعتمد الباحثون على الاختبار النفس – عصبي في تشخيص الإعاقة المعرفية، واستخدمت الدراسة محكات دقة التصنيف والحساسية، والمنطقة الواقعة تحت منحنى (ROC)، واختبار (Press Q) للمقارنة بين الأساليب المستخدمة، وتوصلت نتائج الدراسة إلى أن أسلوب التحليل التمييزي يتمتع بدرجة عالية من الدقة والحساسية وقوة التمييز مقارنة بأسلوب تحليل الانحدار اللوجستي.

كما أجرى إدلر وآخرون (Adler et al., 2011) دراسة للمقارنة بين أسلوبي التحليل التمييزي وتحليل الانحدار اللوجستي في التنبؤ بالإصابة بمرض الجلوكوما وذلك اعتماداً على بعض المنبئات مثل: ضغط العين والتاريخ المرضي والعوامل الوراثية، تم اختيار عينة تشمل عدداً من المصابين وغير المصابين (العاديين)، وتوصلت نتائج الدراسة إلى التشابه في دقة التصنيف بين التحليل التمييزي وتحليل النحدار اللوجستي.

وقام فيرراس وفيجوو Ferreras & Feijoo (2009) بدراسة هدفت إلى التعرف على قدرة بعض المتغيرات في التنبؤ بالإصابة بمرض المياه الزرقاء من خلال بعض المتغيرات المنبئة مثل (العمر، ضغط العين، التاريخ المرضي للحالة، استخدام النظارة، ... وغيرها)، وذلك من خلال بعض البيانات الخاصة بعينتين من العاديين والمصابين بالمرض، باستخدام أسلوبي الانحدار اللوجستي الثنائي، والتحليل التمييزي الخطي، وقد انتهت الدراسة إلى وجود تماثل بين نتائج كل من تحليل الانحدار اللوغاريتمي الثنائي والتحليل التمييزي الخطي.

وفي دراسة أجراها بوهر وآخرون (Pohar et al., 2004) للمقارنة بين أسلوبي تحليل الانحدار اللوجستي والتحليل التمييزي الخطي، استخدمت الدراسة مجموعة من محكات المقارنة للحكم على دقة التنبؤ للأسلوبين، وتوصلت إلى أن القدرة التنبؤية متقاربة بين تحليل الانحدار اللوجستي والتحليل التمييزي إذا تحققت الافتراضات. كما أكدت أن التمييزي الأكثر ملاءمة عند تحقق اعتدالية التوزيع، وإذا لم يتم استيفاء افتراضات التحليل التمييزي الخطي

فليس هناك ما يبرر استخدامه، في حين يعطي تحليل الانحدار اللوجستي نتائج جيدة بغض النظر عن شكل التوزيع. كما يتم الحصول على تقديرات تحليل الانحدار اللوجستي بطريقة دالة الأرجحية العظمى والتي تتميز بمزايا أفضل.

وأجرى لي وكويلي (Lei & Kohly, 2003) دراسة هدفت للمقارنة بين التحليل التمييزي الخطي والانحدار اللوجستي، من خلال مقارنه أخطاء التصنيف في مجموعتين، وقد قارن الباحثان بين طريقتين من طرق التصنيف وهي التحليل التمييزي الخطي والانحدار اللوجستي في ظل ظروف افتراضية مختلفة تحاكيها البيانات المولدة، كحجم العينة، التوازن بين أحجام المجموعات، كما واعتمدا تساوي مصفوفة التباين المشترك أو عدم تساويها بعين الاعتبار، ليقوما بتقصي أثر تلك الظروف كمتغيرات مستقلة على تصنيف الحالات بشكل صحيح في مجموعتين باستخدام ثلاث متنبئات، وقد أظهرت نتائج الدراسة إذا لم يتم أخذ حجم العينة بعين الاعتبار أن أداء الانحدار اللوجستي أفضل من التحليل التمييزي الخطي.

**منهج الدراسة**

اعتمد الباحث في هذه الدراسة على المنهج الوصفي التحليلي.

**مجتمع الدراسة وعينتها**

تكون مجتمع الدراسة من جميع الطلبة الذين قبلوا للدراسة في جامعة القصيم في مسارات السنة التحضيرية المختلفة وتم تسكينهم في تخصصاتهم العلمية والأدبية وانتهوا من دراسة عام أكاديمي في التخصص بعد السنة التحضيرية للعام الجامعي 2021/2020م والبالغ عددهم (١٠٨١) طالباً وطالبة.

وتكونت عينة الدراسة من (١٠٦٨) طالباً وطالبة بجامعة القصيم ممن أنهوا دراسة عام أكاديمي في التخصص بعد السنة التحضيرية وذلك بعد استبعاد الحالات التي يوجد نقص في بياناتها للعام الجامعي 2021/2020م، والجدول (١) يبين توزيع عينة الدراسة حسب متغيرات الدراسة.

**جدول ١**

***توزيع أفراد عينة الدراسة حسب متغيرات الدراسة***

| المتغير | الفئات | التكرار | النسبة المئوية |
|---|---|---|---|
| الجنس | ذكور | ٥٧٥ | ٥٣,٨ |
| | إناث | ٤٩٣ | ٤٦,٢ |
| المجموعة | متعثر | ٨٨ | ٨,٢ |
| | غير متعثر | ٩٨٠ | ٩١,٨ |
| المجموع | | ١٠٦٨ | ١٠٠ |

**أدوات الدراسة**

تم الاعتماد في هذه الدراسة على بيانات الطلبة والمكونة من درجاتهم في: معدل الثانوية العامة، والاختبار التحصيلي، واختبار القدرات العامة، والمعدل التراكمي للسنة التحضيرية، والمعدل التراكمي للسنة الأولى التي تلت السنة التحضيرية، والتي تم الحصول عليها من قبل عمادة القبول والتسجيل في جامعة القصيم.

**متغيرات الدراسة**

- المتغيرات المستقلة (المتنبئة): معدل الثانوية العامة، درجات الاختبار التحصيلي، درجات اختبار القدرات، معدل التخصيص (المعدل التراكمي في السنة التحضيرية).
- المتغير التابع (المتنبأ به): تصنيف الطلبة إلى متعثر وغير متعثر.

**الأساليب الإحصائية**

للإجابة على أسئلة الدراسة تم استخدام الرزمة الإحصائية للعلوم الاجتماعية (SPSS) لأسلوب الانحدار اللوجستي الثنائي والتحليل التمييزي.

**نتائج الدراسة ومناقشتها**

**أولاً: النتائج المتعلقة بسؤال الدراسة الأول والذي ينص على:** " ما أبرز العوامل التي تساهم في التنبؤ بتعثر الطلبة دراسياً باستخدام كل من أسلوب التحليل التمييزي، وأسلوب تحليل الانحدار اللوجستي؟"

### ❖ تطبيق الانحدار اللوجستي التدريجي Stepwise Logistic Regression

للإجابة عن سؤال الدراسة الأول تم استخراج المتغيرات المستقلة الداخلة في التحليل لنموذج الانحدار اللوجستي والتي لها تأثير على تصنيف الطلبة المتعثرين، حيث تم استخدام طريقة الأرجحية العظمى ML في حل النموذج المقترح للحصول على تقديرات ذات كفاءة عالية باستخدام أسلوب الانحدار اللوجستي التدريجي (Stepwise) والجدول (٢) يوضح ذلك.

**جدول ٢**

***المتغيرات الداخلة في التحليل باستخدام تحليل الانحدار اللوجستي التدريجي***

| الخطوات | المتغيرات | المعاملات B | الخطأ المعياري | قيمة إحصائية وولد Wald | درجة الحرية | الدلالة الإحصائية | الدالة الأسية للمعاملات Ex(b) |
|---|---|---|---|---|---|---|---|
| ١ | معدل الثانوية ($X_1$) | ٠,٥٢٨- | ٠,٠٤٥ | ٠,٠٤٥ | ١ | ٠,٠٠ | 0.59 |
| | الثابت | ٤٨,١٠٥ | ٤,٢٤٨ | ٤,٢٤٨ | ١ | ٠,٠٠ | 7.79E+20 |
| ٢ | معدل الثانوية ($X_1$) | ٠,٧١٨- | ٠,١٠٢ | ٠,١٠٢ | ١ | ٠,٠٠ | 0.488 |
| | معدل التخصيص ($X_4$) | ٨,٤٦٦- | ١,١٠١ | ١,١٠١ | ١ | ٠,٠٠ | 0 |
| | الثابت | ٨٧,٩٦٧ | ١١,٤٩ | ١١,٤٩ | ١ | ٠,٠٠ | 1.60E+38 |

| الخطوات | المتغيرات | المعاملات B | الخطأ المعياري | قيمة إحصائية وولد Wald | درجة الحرية | الدلالة الإحصائية | الدالة الأسية للمعاملات Ex(b) |
|---|---|---|---|---|---|---|---|
| ٣ | معدل الثانوية ($x_1$) | ٠,٧١٥- | ٠,١٠١ | ٠,١٠١ | ١ | ٠,٠٠ | 0.489 |
| | التحصيلي ($x_3$) | ٠,٠٦١ | ٠,٠٣ | ٠,٠٣ | ١ | ٠,٠٠ | 1.063 |
| | معدل التخصيص ($x_4$) | ٨,٣٣٧- | ١,١٠٨ | ١,١٠٨ | ١ | ٠,٠٠ | 0 |
| | الثابت | ٨٢,٦٦٩ | ١١,٤٤٣ | ١١,٤٤٣ | ١ | ٠,٠٠ | 7.99E+35 |

١. المتغيرات التي دخلت في الخطوة الأولى: معدل الثانوية

٢. المتغيرات التي دخلت في الخطوة الثانية: معدل التخصيص

٣. المتغيرات التي دخلت في الخطوة الثالثة: التحصيلي

تشير النتائج الموضحة في الجدول رقم (٢) إلى دخول ثلاثة متغيرات في النموذج وهي على التوالي: معدل الثانوية، معدل التخصيص، ودرجة الاختبار التحصيلي، وهي ذات دلالة إحصائية ولها تأثير واضح على تصنيف الطلبة المتعثرين دراسياً، كما يبين الجدول معاملات هذه المتغيرات، وهي لمتغير معدل الثانوية (0.715-) والذي يمكن تفسيره على أن لوغاريتم معامل ترجيح متغير تصنيف الطالب المتعثر دراسياً يقل بمقدار (0.715) كلما زادت معدل الثانوية درجة واحدة مع ضبط أثر المتغيرات المستقلة الأخرى، وكانت لمتغير درجة الاختبار التحصيلي (0.061) وهذا يفسر بأن أن لوغاريتم معامل ترجيح متغير تصنيف الطالب المتعثر دراسياً يزيد مقدر (0.061) كلما زادت درجة متغير الاختبار التحصيلي درجة واحدة، وكانت كذلك لمتغير معدل التخصيص (8.337-) وهذه تفسر على أن لوغاريتم معامل ترجيح متغير تصنيف الطالب المتعثر دراسياً يقل بمقدار (8.337) كلما زاد معدل التخصيص درجة واحدة فقط.

وبالتالي فإن معادلة الانحدار اللوجستي هي:

$$\log\left(\frac{p}{1-p}\right) = 82.699 - 0.715x_1 + 0.061x_3 - 8.337x_4$$

### ❖ تطبيق التحليل التمييزي التدريجي Stepwise Discriminant Analysis

وللإجابة عن سؤال الدراسة الأول تم تطبيق التحليل التمييزي التدريجي Stepwise DA لعرض الخطوات التي مر بها البرنامج من إدخال واستبعاد المتغيرات المستقلة من التحليل، والجدول (3) يبين الخطوات التي مر بها البرنامج وترتيب المتغيرات أو استبعادها من التحليل.

**جدول ٣**

***الخطوات التي مر بها البرنامج لإدخال المتغيرات المستقلة باستخدام التحليل التمييزي التدريجي***

| الخطوة | إدخال المتغير | ولكس لامدا Wilk's Lambda | | | اختبار F | | | |
|---|---|---|---|---|---|---|---|---|
| | | الإحصاء | df1 | df2 | df3 | الإحصاء | df1 | df2 | الدلالة الإحصائية |
| 1 | معدل الثانوية ($X_1$) | ٠,٧٤٨ | ١ | ١ | ١٠٦٦ | ٣٥٩,٤٩ | ١ | ١٠٦٦ | ٠ |
| 2 | معدل التخصيص ($X_4$) | ٠,٦٥٧ | ٢ | ١ | ١٠٦٦ | ٢٧٧,٧٤٣ | ٢ | ١٠٦٥ | ٠ |
| 3 | التحصيلي ($X_3$) | ٠,٦٣٥ | ٣ | ١ | ١٠٦٦ | ٢٠٣,٩٦٤ | ٣ | ١٠٦٤ | ٠ |

أ. الحد الأعلى للخطوات = 8

ب. الحد لمستوى الدلالة لدخول المتغير = 0.05

ج. الحد الأعلى لمستوى الدلالة لـ (F) التي يحذف عندها المتغير من التحليل = 0.10

تشير النتائج في الجدول رقم (٣) إلى دخول ثلاثة متغيرات فقط للنموذج، وهي: معدل الثانوية، ومعدل التخصيص، ودرجة الاختبار التحصيلي، وهي أكثر المتغيرات تصنيفاً للطلبة المتعثرين دراسياً، وقاعدة دخول المتغير في التحليل هي أن المتغير الذي يجعل القيمة الكلية لولكس لامدا Wilk's Lambda نهاية صغرى هو الذي يدخل في التحليل ثم الذي يليه وهكذا؛ أي كلما كانت قيمة ولكس لامدا Wilk's Lambda أقل للمتغير المستقل ودالة إحصائياً كلما كانت أهمية هذا المتغير أقوى في القدرة التصنيفية والتنبؤية، كما تشير النتائج إلى أن قيمة اختبار F لمتغير معدل الثانوية بلغ (359.49) وقيمة ولكس لامدا بلغت (0.748) وهي دالة إحصائياً عند مستوى دلالة (0.01) وهذا يعني أن متغير معدل الثانوية يعتبر أكثر أهمية في النموذج، ثم يليه متغير معدل التخصيص الذي بلغت قيمة اختبار F له (277.743) وقيمة ولكس لامدا بلغت (0.657) وهي دالة إحصائياً، ثم يليه متغير درجة الاختبار التحصيلي التخصيص الذي بلغت قيمة اختبار F له (203.964) وقيمة ولكس لامدا بلغت (0.635) وهي دالة إحصائياً كذلك.

وتم استخراج المعاملات المعيارية للمتغيرات المستخدمة في النموذج والتي لها تأثير على تصنيف الطلبة المتعثرين دراسياً، وهي كما يوضحها الجدول رقم (٤).

**جدول ٤**

***المعاملات المعيارية للمتغيرات الداخلة في النموذج وفق أسلوب التحليل التمييزي التدريجي***

| المتغير | المعامل |
|---|---|
| معدل الثانوية ($X_1$) | ٠,٦٧٢ |
| التحصيلي ($X_3$) | -٠,٠٣٢٤ |
| معدل التخصيص ($X_4$) | ٠,٦٧٣ |

يبين الجدول (٤) المعاملات المعيارية للمتغيرات التي دخلت في النموذج ولها تأثير على تصنيف الطلبة المتعثرين دراسياً، ومن الملاحظ أن متغير التخصيص الأكثر أهمية ومساهمة في القدرة على التمييز

بين الطلبة المتعثرين وغير المتعثرين، حيث بلغت قيمة المعامل له (0.673)، يليه متغير درجة الاختبار التحصيلي، يليه من حيث الأهمية متغير معدل الثانوية الذي بلغت قيمة المعامل له (0.672)، ثم متغير درجة الاختبار التحصيلي والذي بلغ المعامل له (0.324)، ومن هذه المعاملات فإن شكل الدالة التمييزية تأخذ الشكل الآتي:

$$y = 0.672x_1 - 0.324x_3 + 0.673x_4$$

**ثانياً: النتائج المتعلقة بسؤال الدراسة الثاني والذي ينص على:** "ما مدى دقة كل من أسلوبي التحليل التمييزي وتحليل الانحدار اللوجستي في تصنيف الطلبة المتعثرين وغير المتعثرين دراسياً؟"

### ❖ عرض وتحليل نتائج نموذج تحليل الانحدار اللوجستي

للإجابة على سؤال الدراسة الثاني والكشف عن دقة أسلوب تحليل الانحدار اللوجستي في تصنيف الطلبة المتعثرين وغير المتعثرين، فقد تم توليف بيانات المتغيرات الأربعة (معدل الثانوية العامة، درجة الاختبار التحصيل، درجة اختبار القدرات، ومعدل التخصيص) مع المتغير التابع ثنائي القيمة (تصنيف الطالب إلى متعثر وغير متعثر) باستخدام نموذج تحليل الانحدار اللوجستي، وللكشف عن مدى ملائمة نموذج الانحدار اللوجستي وقدرة المتغيرات التي تم إدخالها في النموذج لتصنيف الطلبة المتعثرين وغير المتعثرين، تم إضافة جميع المتغيرات المستقلة للتعرف على أثرها مجتمعة في المتغير التابع لتحديد كفاءة النموذج ككل، حيث قام الباحث باختبار كفاءة النموذج الذي تم توليفه من المتغيرات المستقلة الأربعة وذلك بحساب الدلالة الإحصائية لهذا النموذج باستخدام اختبار مربع كاي، والجدول رقم (٥) يوضح ذلك.

**جدول ٥**

***اختبار الدلالة الإحصائية وقيمة مربع كاي لنموذج الانحدار اللوجستي***

| قيمة مربع كاي | درجات الحرية | الدلالة الإحصائية |
|---|---|---|
| ٤٦٥,٩٨١ | ٤ | ٠,٠٠٠ |

توضح النتائج في الجدول (٥) أن قيمة مربع كاي بلغت (٤٦٥,٩٨١) بدرجات حرية (٤) ومستوى دلالة إحصائية (٠,٠٠٠) مما يدل على النموذج الذي تم توليفه والمتضمن المتغيرات الأربعة (معدل الثانوية العامة، درجة الاختبار التحصيل، درجة اختبار القدرات، ومعدل التخصيص) ذا دلالة إحصائية، وهذا يعني أن النموذج له قدرة تنبؤية وقادر على تصنيف الطلاب إلى متعثرين وغير متعثرين، وأن المتغيرات المستقلة الأربعة لها أهمية ومساهمة وتأثير ذات دلالة إحصائية في تصنيف الطلاب إلى متعثرين وغير متعثرين، حيث تؤكد هذه النتيجة على أن هنالك على الأقل عاملاً واحداً من معاملات النموذج لا يساوي صفراً، وبالتالي فإن هنالك على الأقل متغيراً مستقلاً له تأثير وأهمية ومساهمة في تصنيف الطلاب المتعثرين وغير المتعثرين.

وكذلك قام الباحث بالتحقق من الدلالة العملية لنموذج الانحدار اللوجستي المتضمن المتغيرات الأربعة (معدل الثانوية العامة، درجة الاختبار التحصيل، درجة اختبار القدرات، ومعدل التخصيص)

باستخدام قيم مربع (ر) ($R^2$) لكل من كوكس وسنيل $R^2_{cox-snell}$، وللنايغليكيرك $R^2_{Nagelkerrke}$، والجدول رقم (٦) يوضح ذلك.

**جدول ٦**

***مقاييس الدلالة العملية لنموذج الانحدار اللوجستي المتضمن جميع المتغيرات المستقلة***

| سالب لوغاريتم الأرجحية<br>-2 Log likelihood | مربع (ر) لكوكس وسنيل<br>Snell R Square & Cox | مربع (ر) لنايغليكيرك<br>Nagelkerke R Square |
|---|---|---|
| ١٤١,٨٩٣ | ٠,٣٥٤ | ٠,٨١٥ |

تشير نتائج الجدول (٦) أن قيمة مربع (ر) لكوكس وسنيل بلغت (٠,٣٥٤) وهذا يعني أن نموذج الانحدار اللوجستي والمتضمن المتغيرات المستقلة الأربعة يسهم بنسبة (٠,٣٥٤) تخفيض قيمة لوغاريتم دالة الأرجحية (-2 Log likelihood)، بينما بلغت قيمة مربع (ر) لنايغليكيرك (0.815) وهذا يعني أن نموذج الانحدار اللوجستي والمتضمن المتغيرات المستقلة الأربعة يسهم بنسبة (٠,٨١٥) في تخفيض قيمة لوغاريتم دالة الأرجحية (-2 Log likelihood)، مما يشير إلى أن النموذج له قدرة أفضل على تصنيف الطلاب إلى متعثرين وغير متعثرين دراسياً.

وتم حساب تقديرات معاملات نموذج الانحدار اللوجستي للمتغيرات المستقلة الأربعة، وهي كما يوضحها الجدول رقم (٧).

**جدول ٧**

***تقديرات معالم نموذج الانحدار اللوجستي المتضمن جميع المتغيرات المستقلة***

| المتغيرات | المعاملات<br>B | الخطأ المعياري | قيمة إحصائية وولد<br>Wald | درجة الحرية | الدلالة الإحصائية | الدالة الأسية للمعاملات<br>Ex(b) |
|---|---|---|---|---|---|---|
| معدل الثانوية ($x_1$) | -٠,٧١٨ | ٠,١٠١ | ٥٠,٤٤٢ | ١ | ٠,٠٠٠ | 0.488 |
| القدرات ($x_2$) | ٠,٠٢٦ | ٠,٠٣٨ | ٠,٤٧٨ | ١ | ٠,٤٨٩ | 1.027 |
| التحصيلي ($x_3$) | ٠,٠٦ | ٠,٠٣ | ٣,٨٦ | ١ | ٠,٠٤٩ | 1.061 |
| معدل التخصيص ($x_4$) | -٨,٤٨٦ | ١,١٤٨ | ٥٤,٦٣٩ | ١ | ٠,٠٠٠ | 0.000 |
| الثابت | ٨١,٣٧٢ | ١١,٥٢٧ | ٤٩,٨٣٧ | ١ | ٠,٠٠٠ | 2.18E+35 |

تشير نتائج الجدول (٧) أن قيمة معامل الانحدار لمتغير معدل الثانوية (-٠,٧١٨) وتفسر على أن لوغاريتم معامل ترجيح متغير تصنيف الطالب المتعثر دراسياً يقل بمقدار القيمة (٠,٧١٨) كلما زادت درجات متغير معدل الثانوية درجة واحدة، وذلك بعد ضبط أثر المتغيرات المستقلة الأخرى، وكانت كذلك قيمة معامل الانحدار لمتغير درجة الطالب في اختبار القدرات (٠,٠٢٦) وتفسر على أن لوغاريتم معامل ترجيح متغير تصنيف الطالب المتعثر دراسياً يزيد بمقدار (٠,٠٢٦) كلما زادت درجات اختبار القدرات درجة واحدة، وبلغت قيمة معامل الانحدار لمتغير درجات الاختبار التحصيلي(٠,٠٦٠) وتفسر على أن لوغاريتم معامل ترجيح متغير تصنيف الطالب المتعثر دراسياً يزيد بمقدار (٠,٠٦٠) كلما زادت درجات

الاختبار التحصيلي درجة واحدة، فيما بلغت قيمة معامل الانحدار لمتغير معدل التخصيص (-٨,٤٨٦) وتفسر على أن لوغاريتم معامل ترجيح متغير تصنيف الطالب المتعثر دراسياً يقل بمقدار القيمة (٨,٤٨٦) كلما زادت درجات متغير معدل التخصيص درجة واحدة. ويلاحظ من التفسير السابق أنه كلما ارتفعت قيمة معامل الانحدار دل ذلك على أهمية المتغير المستقل في التنبؤ بالمتغير التابع، وكذلك كلما اقترب من الصفر دل ذلك على عدم أهمية المتغير المستقل في التنبؤ بالمتغير التابع. ويمكن كتابة معاجلة الانحدار اللوجستي وفق النتائج السابقة كالآتي:

$$\log\left(\frac{p}{1-p}\right) = 81.372 - 0.718x_1 + 0.026x_2 + 0.060x_3 - 8.486x_4$$

ولاختبار دلالة معالم نموذج الانحدار اللوجستي من القدرة التنبؤية للمتغيرات المستقلة الأربعة (معدل الثانوية العامة، درجة الاختبار التحصيل، درجة اختبار القدرات، ومعدل التخصيص) تم الاعتماد على إحصائية وولد (Wald Statistic)، والموضحة نتائجها في الجدول رقم (٧) السابق والذي يوضح هذه القيم والدلالة الإحصائية لكل متغير من المتغيرات المستقلة.

حيث تشير النتائج في الجدول (٧) إلى أن قيمة والد لمتغير معدل الثانوية بلغ (٥٠,٤٤٢) وهي ذات دلالة إحصائية، وهذا يعني أن قيمة معامل الانحدار اللوجستي لمتغير معدل الثانوية تختلف عن الصفر، وبالتالي فإن متغير معدل الثانوية له دلالة إحصائية في القدرة التنبؤية على احتمال تصنيف الطالب المتعثر دراسياً، كما كانت قيمة اختبار وولد لمتغير اختبار القدرات عند درجة حرية واحدة بلغ (٠,٤٧٨) وهي غير دالة إحصائياً، وهذا يعني أن قيمة معامل الانحدار اللوجستي لمتغير القدرات لا تختلف عن الصفر، وأن متغير اختبار القدرات ليس له دلالة إحصائية في القدرة التنبؤية على احتمال تصنيف الطالب المتعثر دراسياً، وبلغت قيمة اختبار وولد لمتغير درجة الاختبار التحصيلي (٣,٨٦٠) وهي ذات دلالة إحصائية، وهذا يعني قيمة معامل الانحدار اللوجستي لمتغير درجات الاختبار التحصيلي تختلف عن الصفر، وبالتالي فإن متغير درجات الاختبار التحصيلي له دلالة إحصائية في القدرة التنبؤية على احتمال تصنيف الطالب المتعثر دراسياً، في حين بلغت قيمة إحصائية وولد لمتغير معدل التخصيص (٥٤,٦٣٩) وهي ذات دلالة إحصائية، وهذا يعني قيمة معامل الانحدار اللوجستي لمتغير معدل التخصيص تختلف عن الصفر، وبالتالي فإن متغير معدل التخصيص له دلالة إحصائية في القدرة التنبؤية على احتمال تصنيف الطالب المتعثر دراسياً.

يتضح من التفسير السابق أن معاملات الانحدار الخاصة بالمتغيرات (معدل الثانوية، درجات الاختبار التحصيلي، معدل التخصيص) كانت ذات دلالة إحصائية، أي لها تأثير دال إحصائياً على احتمال تصنيف الطالب المتعثر دراسياً. أما متغير اختبار القدرات ليس له تأثير دال إحصائياً على احتمال تصنيف الطالب المتعثر دراسياً.

وللحكم على القدرة التنبؤية وفحص الدلالة لكل متغير من المتغيرات المستقلة الأربعة، تم استخدام اختبار الدرجة Score Test والجدول رقم (٨) يبين ذلك.

**جدول ٨**

***إحصاء الدرجة Score للمتغيرات المستقلة (معدل الثانوية، درجات الاختبار التحصيلي، معدل التخصيص) وفق نموذج الانحدار اللوجستي***

| المتغيرات | إحصاء الدرجة score | درجات الحرية | الدلالة الإحصائية |
|---|---|---|---|
| معدل الثانوية ($x_1$) | ٢٦٩,٣٤ | ١ | ٠,٠٠٠ |
| القدرات ($x_2$) | ٢١,١٤ | ١ | ٠,٠٠٠ |
| التحصيلي ($x_3$) | ٠,٤٦ | ١ | ٠,٤٣٨ |
| معدل التخصيص ($x_4$) | ٢٥٣,٣٦ | ١ | ٠,٠٠٠ |

يتبين من نتائج الجدول (٨) أن إحصائية الدرجة للمتغيرات الأربعة (معدل الثانوية العامة، درجة الاختبار التحصيل، درجة اختبار القدرات، ومعدل التخصيص) دالة إحصائياً، وهذا يعني أن معامل كل متغير من المتغيرات الأربعة في نموذج الانحدار اللوجستي يختلف عن الصفر، وبالتالي فإن كل متغير له دلالة إحصائية في تصنيف المتغير التابع؛ وبذلك فإن النتيجة تدل على أن المتغيرات الأربعة تتمتع بقدرة تنبؤية وتفسيرية للمتغير التابع تختلف عن الصدفة.

وللتأكد من وجود فروق ذات دلالة إحصائية بين القيم المشاهدة والقيم المتوقعة للبيانات، تم استخدام اختبار هوزمر وليمشو Hosmer & Lemeshw لجودة المطابقة كما هو موضح في الجدول رقم (٩).

**جدول ٩**

***اختبار هوسمر وليمشو Hosmer & Lemeshw لجودة المطابقة***

| قيمة مربع كاي | درجات الحرية | الدلالة الإحصائية |
|---|---|---|
| ٠,٦٧٦ | ٨ | ١,٠٠ |

تشير النتائج في الجدول أعلاه أن قيمة اختبار مربع كاي تساوي (٠,٦٧٦) عند درجة حرية تساوي (٨) ودلالة إحصائية تساوي (1.00) وهي أكبر من (0.05)، وبالتالي فهي دالة إحصائياً، ويدل ذلك على قبول الفرض الصفري في هذه الحالة القائل بأن البيانات المتوقعة بالنموذج المستخدم تطابق البيانات المشاهدة، وهذا تعزيز للنموذج المستخدم.

ولغرض تقييم مدى نجاح نموذج الانحدار اللوجستي في تصنيف الطلبة المتعثرين، تم حساب دقة التصنيف لدى العينة على فئتي المتغير التابع (متعثر، غير متعثر)، ويوضح الجدول رقم (١٠) النسبة المئوية للتصنيف المشاهد والمتوقع.

**جدول ١٠**

***النسبة المئوية للتصنيف الصحيح المشاهد والمتوقع للطلبة المتعثرين وغير المتعثرين دراسياً في نموذج الانحدار اللوجستي***

| التصنيف المشاهد | التصنيف المتوقع | | النسبة المئوية للتصنيف الصحيح |
|---|---|---|---|
| | متعثر | غير متعثر | |
| متعثر | ٨٠ | ٨ | ٩٠,٩ |
| غير متعثر | ١٥ | ٩٦٥ | ٩٨,٥ |
| النسبة الكلية | | | ٩٧,٨ |

يلاحظ من الجدول أعلاه أن جودة التصنيف ممتازة لنموذج الانحدار اللوجستي، حيث بلغت نسبة التصنيف الصحيح (٩٧,٨)، وهذا يعني أن جميع الطلبة تم تصنيفهم من قبل النموذج كما هي عليه فعلاً، إلا (١٥) طلاب صُنّفوا على أنهم طلبة متعثرين في البيانات المشاهدة إلا أنه تم تصنيفهم في النموذج على أنهم طلبة غير متعثرين، و(٨) طلاب صُنّفوا على أنهم طلبة غير متعثرين في البيانات المشاهدة إلا أنه تم تصنيفهم في النموذج على أنهم طلبة متعثرين.

### ❖ عرض وتحليل نتائج نموذج التحليل التمييزي

للإجابة على سؤال الدراسة الثاني والكشف عن دقة نموذج التحليل التمييزي في تصنيف الطلبة المتعثرين وغير المتعثرين دراسياً، في البداية تم التأكد من تحقق افتراضات النموذج، حيث من الملاحظ وجود 1068حالة، وكذلك المتغيرات المستقلة الأربعة هي متغيرات كمية، وبالتالي وفقاً لنظرية النهاية المركزية فإن البيانات تتبع التوزيع الاعتدالي، ولمزيد من التأكيد تم استخدام اختبار كولموغوروف-سميرنوف (Kolmogrov-Smirnov) لاختبار اعتدالية التوزيع، والجدول رقم (١١) يبين تلك النتائج.

**جدول ١١**

***اختبار كولموغوروف-سميرنوف (Kolmogrov-Smirnov) لاختبار اعتدالية التوزيع للمتغيرات المستقلة (معدل الثانوية، درجات الاختبار التحصيلي، معدل التخصيص)***

| المتغيرات | قيمة الإحصائي | درجة الحرية | الدلالة الإحصائية |
|---|---|---|---|
| معدل الثانوية ($X_1$) | ٠,٢٨ | ١٠٦٨ | ٠,١٩١ |
| القدرات ($X_2$) | ٠,٠١٧ | ١٠٦٨ | ٠,٢ |
| التحصيلي ($X_3$) | ٠,٠١٩ | ١٠٦٨ | ٠,٢ |
| معدل التخصيص ($X_4$) | ٠,٠٢٥ | ١٠٦٨ | ٠,١١ |

تشير نتائج الجدول أعلاه أن جميع المتغيرات المستقلة تتبع التوزيع الاعتدالي وذلك اعتمادا على قيمة إحصائي كولموغوروف-سميرنوف والدلالة الإحصائية له، مما يعني تحقق أحد افتراضات التحليل التمييزي، وبما أن جميع المتغيرات المستقلة هي متغيرات كمية ونظراً لكون حجم العينة قد تجاوزت (٣٠) مشاهدة، حيث تم استخدام (١٠٦٨) مشاهدة، وبالتالي فإنه وفقاً لنظرية النهاية المركزية فإن البيانات تتبع التوزيع الاعتدالي.

وللتحقق من افتراض تجانس التباين بين مجموعتي الطلبة المتعثرين وغير المتعثرين دراسياً، تم حساب قيمة (Box's M)، والنتائج كما هي موضحة في الجدول رقم (١٢).

**جدول ١٢**

***اختبار Box's M لتجانس التباين بين مجموعتي الطلبة المتعثرين وغير المتعثرين دراسياً***

| | | |
|---|---|---|
| | قيمة اختبار Box's M | ١٦٨,٣٧٧ |
| اختبار F | القيمة | ١٦,٥٥٧ |
| | درجات الحرية 1 | ١٠ |
| | درجات الحرية 2 | ٩٨٧٩٥,٢ |
| | الدلالة الإحصائية | ٠,٠٦٧ |

يتضح من نتائج الجدول أعلاه أن قيمة Box's M بلغت (168.377) وهي دالة إحصائياً، مما يعني عدم تحقق افتراض تجانس التباين والتغاير، حيث يعتبر هذا الاختبار حساس جداً في لاختبار فرضية التجانس للمجتمع، حيث أنه في الحالات التي يكون عدد أفراد عينتها كبير تظهر الاختلافات البسيطة فروقاً في النتائج، وتشير الدراسات إلى أن نموذج التحليل التمييزي يمكن أن يصمد في مواجهة هذا الافتراض.

وللكشف عن دقة نموذج التحليل التمييزي في تصنيف الطلبة المتعثرين دراسياً تم استخدام جميع المتغيرات المستقلة الأربعة (معدل الثانوية، درجات الاختبار التحصيلي، معدل التخصيص) مع المتغير التابع لتصنيف الطلبة المتعثرين دراسياً باستخدام نموذج التحليل التمييزي، ولتحديد العوامل التمييزية والأهمية النسبية لكل متغير مستقل على حده ومدى تأثيره في بناء الدالة التمييزية؛ تم حساب قيمة ولكس لامدا Wilk's Lambda واختبار النسبة الفائية للمتغيرات المستقلة الأربعة كما هو موضح في الجدول رقم (١٣).

**جدول ١٣**

***اختبار ولكس لامدا Wilk's Lambda والنسبة الفائية (F) لمتغيرات الدالة التمييزية المستقلة (معدل الثانوية، درجات الاختبار التحصيلي، معدل التخصيص)***

| المتغيرات | ولكس لامدا | اختبار F | درجات الحرية 1 | درجات الحرية 2 | الدلالة الإحصائية |
|---|---|---|---|---|---|
| معدل الثانوية ($x_1$) | ٠,٧٤٨ | ٣٥٩,٤٩ | ١ | ١٠٦٦ | ٠,٠٠٠ |
| القدرات ($x_2$) | ٠,٩٨ | ٢١,٥٢١ | ١ | ١٠٦٦ | ٠,٠٠٠ |
| التحصيلي ($x_3$) | ٩٩٢ | ٠,٤٥٨ | ١ | ١٠٦٦ | ٠,٠٤٧ |
| معدل التخصيص ($x_4$) | ٠,٧٦٣ | ٣٣١,٥٣٩ | ١ | ١٠٦٦ | ٠,٠٠٠ |

تشير النتائج في الجدول أعلاه إلى أن قيمة اختبار ولكس لامدا Wilk's Lambda للمتغيرات الأربعة المستقلة (معدل الثانوية، درجات اختبار القدرات، درجات الاختبار التحصيلي، معدل التخصيص) هي قيم دالة إحصائياً عند مستوى دلالة (0.01)، ويدل ذلك على أن جميع المتغيرات المستقلة الأربعة لها دلالة إحصائية في القدرة التنبؤية، وأن كل متغير منها له تأثير واهمية في تكوين وبناء الدالة التمييزية.

وكلما كانت قيمة ولكس لامدا Wilk's Lambda أقل للمتغير المستقل ودالة إحصائياً كلما كانت أهمية هذا المتغير أقوى في القدرة التصنيفية والتنبؤية.

كما يُفسر اختبار (F) الأهمية النسبية للمتغيرات المستقلة المكونة لنموذج التحليل التمييزي، وكلما كانت قيمة اختبار (F) مرتفعة ودالة إحصائياً كلما للمتغير المستقل أهمية، والنتائج في الجدول أعلاه تشير جميع المتغيرات دالة إحصائياً بالنسبة لقيمة اختبار (F) للمتغيرات ويمكن ترتيبها حسب أهميتها النسبية استناداً لتلك النتائج كالآتي:متغير معدل الثانوية، يليه معدل التخصيص، يليه اختبار القدرات، يليه الاختبار التحصيلي، وهذا ما اتفقت عليه قيم ولكس لامدا Wilk's Lambda في تحديد المتغيرات المستقلة ذات الأهمية في النموذج.

ولتحديد الدالة التمييزية وكذلك النسبة المئوية من التباين التي تُعزى إلى الدالة التمييزية، يوضح الجدول رقم (١٤) قيمة الجذر الكامن ونسبة التباين المفسر بين فئتي المتغير التابع (متعثر، وغير متعثر) والارتباط القانوني.

**جدول ١٤**

***قيمة الجذر الكامن والتباين المفسر ومعامل الارتباط القانوني للدالة التمييزية***

| الدالة | قيمة الجذر الكامن | نسبة التباين % | الارتباط القانوني |
|---|---|---|---|
| ١ | ٠,٥٨ | ١٠٠ | ٠,٦٠٦ |

يلاحظ من نتائج الجدول (13) أن قيمة الجذر الكامن بلغت (0580)، وهذه القيمة تشير إلى نسبة التباين المفسر بين مجموعتي الطلبة المتعثرين وغير المتعثرين والتي تعود إلى الفروق بينهما في دالة التمييز الوحيدة، ونجد أن نسبة التباين المفسر بلغ (100%) وسبب ذلك يعود إلى وجود دالة تمييزية واحدة في النموذج، وبلغت قيمة معامل الارتباط القانوني (0.606) وتعني أن النسبة المئوية في المتغير التابع الذي تم تمييزه بالمتغيرات المستقلة الأربعة بلغ (60.6%).

وتم ختبار الدالة التمييزية وقدرتها التنبؤية باستخدام اختبار ولكس لامدا Wilk's Lambda، والجدول رقم (١٥) يوضح ذلك.

**جدول ١٥**

***اختبار ولكس لامدا Wilk's Lambda للدالة التمييزية***

| الدالة | ولكس لامدا | مربع كاي | درجة الحرية | الدلالة الإحصائية |
|---|---|---|---|---|
| ١ | ٠,٦٣٣ | ٤٨٦,٦١٤ | ٤ | ٠,٠٠٠ |

يتضح من الجدول (١٥) أن قيمة ولكس لامدا Wilk's Lambda المستخدم لاختبار الدالة التمييزية بلغ (0.633) وهي دالة إحصائياً، وكذلك قيمة مربع كاي (486.614) بدجات حرية (4) دالة إحصائياً عند مستوى الدلالة (0.01)، يدل ذلك على أن الدالة التمييزية معنوية في القدرة على

تمييز بين الطلبة المتعثرين وغير المتعثرين دراسياً، وبالتالي يعتمد على هذه الدالة في تمييز الطلبة الجدد على المتغير التابع (متعثر، غير متعثر).

ولتحديد مساهمة كل متغير من المتغيرات المستقلة الأربعة في التمييز بين الطلبة المتعثرين وغير المتعثرين دراسياً، تم فحص المعاملات المعيارية للمتغيرات في الدالة التمييزية، والجدول رقم (١٦) يوضح قيم هذه المعاملات.

**جدول ١٦**

***المعاملات المعيارية للمتغيرات المستقلة (معدل الثانوية، درجات الاختبار التحصيلي، معدل التخصيص)***

| المتغيرات | قيم معامل الدالة التمييزية |
|---|---|
| معدل الثانوية ($X_1$) | ٠,٦٨٣ |
| القدرات ($X_2$) | -٠,٠٩٦ |
| التحصيلي ($X_3$) | -٠,٣١٤ |
| معدل التخصيص ($X_4$) | ٠,٦٩ |

تُظهر نتائج الجدول (١٦) مساهمة كل متغير في الدالة التمييزية، حيث يلاحظ أن متغير معدل التخصيص الأكثر أهمية ومساهمة في القدرة على التمييز بين الطلبة المتعثرين وغير المتعثرين دراسياً، حيث بلغت قيمة المعامل له (0.690)، يليه متغير معدل الثانوية حيث بلغت قيمة المعامل له (0.683)، ثم متغير الاختبار التحصيلي حيث بلغت قيمة المعامل له (0.314-)، وأخيراً متغير اختبار القدرات حيث بلغت قيمة المعامل له (0.096-). ومن هذه البيانات فإن الدالة التمييزية تصبح كالآتي:

$$Y = 0.683x_1 - 0.096x_2 - 0.314x_3 + 0.690x_4$$

لاختبار قدرة الدالة على التمييز والتصنيف بين فئتي الطلبة المتعثرين وغير المتعثرين دراسياً تم حساب دقة التصنيف على فئتي المتغير التابع (متعثر، غير متعثر)، ويوضح الجدول رقم (١٧) النسبة المئوية للتصنيف المشاهد والمتوقع.

**جدول ١٧**

***النسبة المئوية للتصنيف الصحيح المشاهد والمتوقع للطلبة المتعثرين وغير المتعثرين دراسياً في نموذج التحليل التمييزي***

| التصنيف المشاهد | التصنيف المتوقع | | النسبة المئوية للتصنيف الصحيح |
|---|---|---|---|
| | متعثر | غير متعثر | |
| متعثر | ٨٦ | ٢ | ٩٧,٧ |
| غير متعثر | ١٠٤ | ٨٧٦ | ٨٩,٤ |
| نسبة التصنيف الصحيح للدالة | | | ٩٠,٠ |

يتضح من نتائج الجدول (١٧) أن نسبة المتنبأ به من أفراد عينة الدراسة من الطلبة المتعثرين دراسياً والذي تم التنبؤ بعضويتهم في مجموعة الطلبة المتعثرين بلغت (97.7%)، في حين بلغت لأفراد عينة الدراسة من الطلبة غير المتعثرين والذي تم التنبؤ بعضويتهم في مجموعة الطلبة غير المتعثرين بلغت (89.4%)، أما إجمالي النسبة المئوية للتصنيف الصحيح فقد بلغت (90%).

**ثالثاً: النتائج المتعلقة بسؤال الدراسة الثالث والذي ينص على:** "أي الأسلوبين المستخدمين: التحليل التمييزي، الانحدار اللوجستي هو الأكثر دقة وفاعلية في تصنيف الطلبة المتعثرين دراسياً؟"

للإجابة عن سؤال الدراسة الثالث فقد تم المقارنة بين نموذجي التحليل التمييزي والانحدار اللوجستي في تصنيف الطلبة المتعثرين دراسياً، حيث تم في البداية بين جداول التصنيف لكلا النموذجين والتي توضح مدى الدقة في تصنيف الطلبة المتعثرين دراسياً، ويعرض الجدول رقم (١٨) ملخصاً لقيم حساسية التصنيف ودقة التصنيف ونسبة التصنيف الصحيحة الكلية لكل من نموذج الانحدار اللوجستي والتحليل التمييزي.

**جدول ١٨**

***قيم حساسية التصنيف ودقة التصنيف ونسبة التصنيف الصحيحة الكلية لنموذجي الانحدار اللوجستي والتحليل التمييزي***

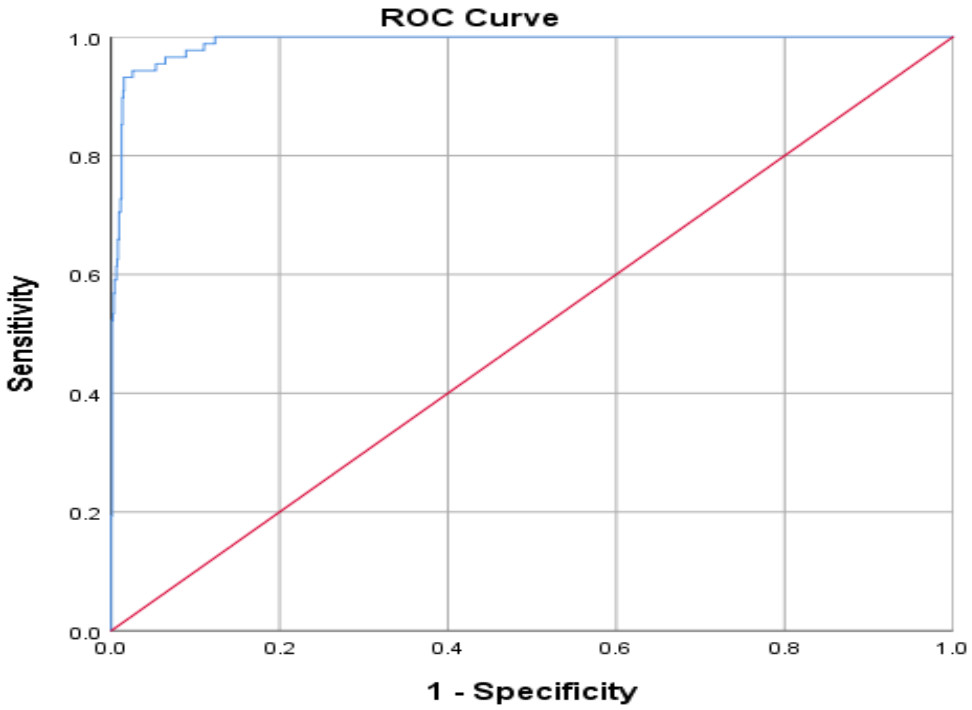
| النموذج | حساسية التصنيف | دقة التصنيف | النسبة المئوية للتصنيف الصحيح |
|---|---|---|---|
| الانحدار اللوجستي | ٩٠,٩ | ٩٨,٥ | ٩٧,٨ |
| التحليل التمييزي | ٩٧,٧ | ٨٩,٤ | ٩٠ |

يتبين من النتائج في الجدول (١٨) أن قيمة حساسية التصنيف لنموذج تحليل الانحدار اللوجستي بلغت (90.9)، بينما في نموذج التحليل التمييزي بلغت (97.7) وبفارق (6.8) لصالح التحليل التمييزي، ونجد أن دقة التصنيف عند استخدام نموذج تحليل الانحدار اللوجستي بلغت (98.5)، في حين بلغت عند استخدام التحليل التمييزي (89.4) وبفارق (9.1) لصالح تحليل الانحدار اللوجستي، كما نجد قيمة التصنيف الصحيح عند استخدام تحليل الانحدار اللوجستي بلغت (97.8)، في حين بلغت عند استخدام نموذج التحليل التمييزي (90.0) وبفارق (7.8) لصالح تحليل الانحدار اللوجستي، وهذا مؤشر على أفضلية نموذج تحليل الانحدار بحسب معيار دقة التصنيف.

وفيما يتعلق بالمقارنة بين نموذج تحليل الانحدار ونموذج التحليل التمييزي من خلال المساحة أسفل المنحنى (ROC)، ففي البداية تم حساب المساحة أسفل المنحنى (ROC) لنموذج الانحدار اللوجستي من أجل قياس دقة تمييز النموذج للطلبة المتعثرين وغير المتعثرين دراسياً، والشكل رقم (٢) يوضح هذا المنحنى.

**الشكل ٢**

***منحنى ROC لنموذج تحليل الانحدار اللوجستي***




يوضح الشكل (2) أن نموذج تحليل الانحدار اللوجستي يعمل على تصنيف البيانات أفضل من عمل عامل الصدفة، نلاحظ أن المنحنى يبتعد عن قطر الصدفة (الخط باللون الأحمر) والمحصور تحته (50%) من المساحة ليعطي مساحة أكبر، ونلاحظ أيضاً أن الخط الأزرق يكاد يشكل زاوية قائمة وهي الحالة المثلى التي يكون فيها التمييز كاملاً؛ أي 100%. والجدول رقم (١٩) يوضح قيمة المساحة الواقعة تحت منحنى ROC.

**جدول ١٩**

***المساحة تحت منحنى ROC لنموذج تحليل الانحدار اللوجستي***

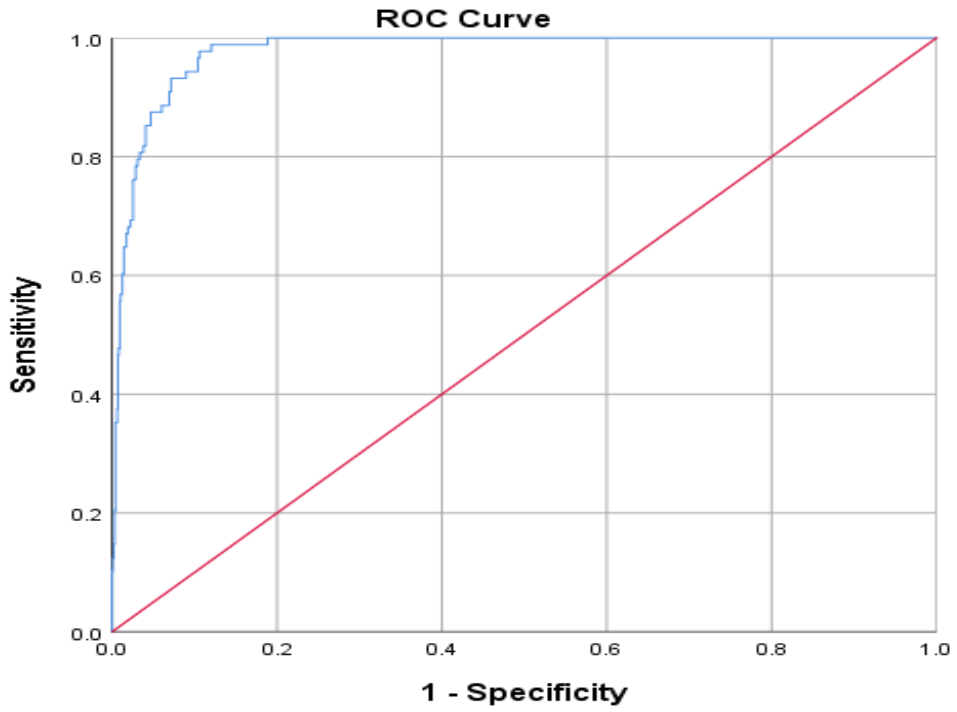
| المساحة | الخطأ المعياري | الدلالة الإحصائية | مستوى الثقة 95% | |
|---|---|---|---|---|
| | | | الحد الأدنى | الحد الأعلى |
| ٠,٩٩ | ٠,٠٠٣ | ٠,٠٠٠ | ٠,٩٨٥ | ٠,٩٩٦ |

حيث يتضح من الجدول (١٩) أن قيمة المساحة تحت المنحنى بلغت (0.99) عند مستوى دلالة أقل من (0.01)، وهذا يؤكد رفض الفرضية الصفرية التي تنص على أن المساحة تحت المنحنى الناتجة من توفيق البيانات باستخدام نموذج تحليل الانحدار اللوجستي تساوي (0.50)، وبالتالي فإن نموذج الانحدار اللوجستي يساعد على التنبؤ بتصنيف حالات المتغير التابع أكثر مما تفعله الصدفة، وتؤكد قيمة المساحة الواقعة تحت المنحنى بأن دقة تمييز نموذج الانحدار اللوجستي للطلبة المتعثرين دراسياً من الطلبة غير المتعثرين ممتازة.

وتم حساب قيمة المساحة تحت منحنى (ROC) لنموذج التحليل التمييزي، والشكل (٣) يوضح هذا المنحنى.

**الشكل ٣**

***منحنى ROC لنموذج التحليل التمييزي***




ونلاحظ من الشكل (٣) أن الخط الأزرق هنا كذلك يكاد يشكل زاوية قائمة وهي الحالة المثلى التي يكون فيها التمييز كاملاً؛ أي 100%. والجدول رقم (٢٠) يوضح قيمة المساحة الواقعة تحت منحنى ROC.

**جدول ٢٠**

***المساحة تحت منحنى ROC لنموذج التحليل التمييزي***

| المساحة | الخطأ المعياري | الدلالة الإحصائية | مستوى الثقة 95% | |
|---|---|---|---|---|
| | | | الحد الأدنى | الحد الأعلى |
| ٠,٩٧٧ | ٠,٠٠٤ | ٠,٠٠٠ | ٠,٩٦٩ | ٠,٩٨٦ |

يتضح من الجدول (٢٠) أن قيمة المساحة تحت المنحنى بلغت (0.977) عند مستوى دلالة أقل من (0.01)، وتؤكد قيمة المساحة الواقعة تحت المنحنى بأن دقة تمييز نموذج التحليل التمييزي للطلبة المتعثرين دراسياً من الطلبة غير المتعثرين ممتازة.

وللمقارنة بين نموذج الانحدار اللوجستي ونموذج التحليل التمييزي من خلال النتائج التي تم التوصل إليه باستخدام المساحة تحت منحنى Roc نلاحظ وجود تفاوت بين المساحات، والجدول رقم (٢١) يوضح ذلك.

**جدول ٢١**

***مقارنة بين المساحة تحت منحنى ROC لنموذجي تحليل الانحدار اللوجستي والتحليل التمييزي***

| النموذج | المساحة تحت منحنى ROC |
|---|---|
| الانحدار اللوجستي | ٠,٩٩ |
| التحليل التمييزي | ٠,٩٧٧ |

يتضح من خلال الجدول (٢١) أن هناك تفاوت بسيط بين قيم المساحة تحت منحنى ROC لنموذج الانحدار اللوجستي والمساحة تحت منحنى ROC لنموذج التحليل التمييزي وبفارق (0.13) لصالح نموذج الانحدار اللوجستي، وبالتالي يتفوق نموذج الانحدار اللوجستي على نموذج التحليل التمييزي بحسب هذا المعيار.

**مناقشة النتائج**

**أولاً: مناقشة النتائج المتعلقة بسؤال الدراسة الأول والذي ينص على:** " ما أبرز العوامل التي تساهم في التنبؤ بتعثر الطلبة دراسياً باستخدام كل من أسلوب التحليل التمييزي، وأسلوب تحليل الانحدار اللوجستي؟"

أظهرت نتائج الدراسة دخول ثلاثة متغيرات فقط داخل النموذج، وترتيب هذه المتغيرات المستقلة المؤثرة في تصنيف الطلبة حسب أهميتها النسبية في نموذج الانحدار اللوجستي هي (معدل الثانوية، معدل التخصيص، درجة الاختبار التحصيلي)؛ أي أن هذه المتغيرات ذات دلالة إحصائية ولها تأثير على تصنيف الطلبة المتعثرين دراسياً. وتتفق هذه النتيجة مع ما توصلت إليه دراسة جريجوري وآخرون (Gregori et al., 2011) في تحديد العوامل التنبؤوية بموت مرض السكر والتي أكدت أن أسلوب الانحدار اللوجستي كان أكثر دقة في تحديد العوامل المنبئة بموت مرض السكر، ودراسة ميهالوفيتش (Mihalovic, 20016) لتحديد أهم المتنبئات الدالة التي تعيق فشل الشركات، والتي أكدت أن النموذج الذي يعتمد على الانحدار اللوجستي يتفوق في دقة التصنيف.

ويمكن تفسير هذه النتيجة إلى أن أسلوب الانحدار اللوجستي يتميز عن نموذج التحليل التمييزي في هذه الدراسة أنه يبين اشارات معاملات النموذج المقدر والتي تبين العلاقة بين كل ستغير مستقل والمتغير التابع، والتي منها يتحدد مدى معاملات النموذج المقدر، والتي تبين العلاقة بين كل متغير مستقل والمتغير التابع، والتي منها يتحدد مدى مطابقة النموذج المقدر للبيانات المشاهدة.

ويمكن تفسير نتيجة أهمية المتغيرات المستقلة التي تم إدخالها في نموذج الانحدار اللوجستي معدل الثانوية ومعدل التخصيص ودرجة الاختبار التحصيلي والتي اعتبرت أنها ذات دلالة إحصائية ولها تأثير على تصنيف الطلبة المتعثرين دراسياً إلى أن أداء الطلبة المتعثرين دراسياً في هذه المتغيرات أقل من أداء الطلبة غير المتعثرين دراسياً ولذلك قد يحصلون على درجات منخفضة في هذه المتغيرات.

وتوصلت الدراسة باستخدام التحليل التمييزي إلى استخلاص ثلاثة متغيرات من بين أربعة متغيرات كأحسن تجمع خطي لها القدرة على التمييز بين الطلبة المتعثرين وغير المتعثرين دراسياً، وكانت أهم المتغيرات حسب ترتيبها في الإدخال: معدل الثانوية، ومعدل التخصيص، ودرجة الاختبار التحصيلي. وهذه النتيجة تتفق مع ما توصلت اليه في دراسة قام بها ليدي آني وجاستون (Lydi-Anne & Gaston, 2011) في تحديد العوامل المنبئة بعدم تعاطي الكحول والعمر للحامل أثناء فترة الحمل وأفضلية أسلوب تحليل الدالة التمييزية في تحديد هذه العوامل.

ويمكن تفسير هذه النتيجة إلى أن نموذج التحليل التمييزي يتميز عن نموذج الانحدار اللوجستي في الدراسة بأنها تستبعد المتغير أو المتغيرات التي ليس لها تأثير معنوي على المتغير التابع، لذلك تم استبعاد متغير اختبار القدرات العامة، وكان ترتيب متغير الثانوية العامة أهم المتغيرات المساهمة في تصنيف الطلبة المتعثرين دراسياً، ويمكن تفسير هذه النتيجة كون أن انخفاض الدرجات التي يحصل عليها الطلبة في الثانوية العامة تؤدي إلى انخفاض تصنيفهم إلى طلبة متعثرين، وهذا منطقي لأن أداء الطلبة المتعثرين في معدل الثانوية أقل من الطلبة غير المتعثرين (النعيمي، ٢٠١٨).

**ثانياً: النتائج المتعلقة بسؤال الدراسة الثاني والذي ينص على:** " ما مدى دقة كل من أسلوبي التحليل التمييزي وتحليل الانحدار اللوجستي في تصنيف الطلبة المتعثرين وغير المتعثرين دراسياً؟"

أظهرت نتائج الدراسة إلى أن النموذج الإحصائي والمتضمن المتغيرات المستقلة (معدل الثانوية، القدرات، درجة الاختبار التحصيلي، معدل التخصيص) ملائم للتنبؤ، وأن المتغيرات المستقلة لها أهمية وتأثير ومساهمة ذات دلالة إحصائية في تصنيف الطلبة المتعثرين وغير المتعثرين دراسياً، وأن الاختلافات بين القيم المشاهدة والقيم المتوقعة كانت بسيطة جداً وهذا ما أظهرته نتيجة قيمة اختبار مربع كاي لهوزمر وليمشو (Hosmer and Lemeshow) بأن البيانات المتوقعة وفق نموذج الانحدار اللوجستي تطابق البيانات المشاهدة، كما دلت نتائج اختبار والد Wald وإحصاءات الدرجة Score في الانحدار اللوجستي على أهمية متغير معدل التخصيص في القدرة التمييزية والتصنيفية، ثم متغير معدل الثانوية، ومن ثم المتغير درجة الاختبار التحصيلي، ويليها متغير درجة اختبار القدرات. كما أشارت النتائج إلى كفاءة وجودة التصنيف لنموذج الانحدار اللوجستي حيث بلغت نسبة لتصنيف الصحيح له (97.8%) وبالتالي يعتبر هذا النموذج ناجح وملائم لتصنيف الطلبة المتعثرين وغير المتعثرين دراسياً. واتفقت هذه النتائج مع دراسة كل من برزنجي (Barznji. 2018) والتي أظهرت أن تحليل الانحدار اللوجستي لديه أفضل تنبؤ لعوامل الخطر لمرض السكري. ودراسة إبراهيم زاده وآخرون (Ebrahimzadeh et al., 2015) والتي توصلت إلى أن دقة التصنيف باستخدام أسلوب نموذج الانحدار اللوجستي كانت أفضل في تصنيف المشاهدات. ودراسة جريجوري وآخرون (Gregori et al., 2011) والتي بينت أن أسلوب الانحدار اللوجستي كان أكثر دقة في تحديد العوامل المنبئة بموت مرض السكر.

ويمكن تفسير هذه النتيجة إلى استخدام نموذج تحليل الانحدار اللوجستي في تقدير معالم النموذج طرقة الأرجحية العظمى (ML) وبالتالي تتوافر في هذه التقديرات شروط الجودة، وكذلك سهولة العمليات الحسابية التي يمتاز بها نموذج الانحدار اللوجستي عند صياغة النموذج مقارنة بالنماذج الأخرى.

فيما أشارت نتائج الدراسة إلى أن الدالة التمييزية لنموذج التحليل التمييزي ذات معنوية ولها قدرة على التمييز بين الطلبة المتعثرين وغير المتعثرين دراسياً، وبذلك يمكن الاعتماد عليها في تمييز الطلبة الجدد بالنسبة للمتغير التابع (متعثر، غير متعثر)، فقد بينت نتائج قيمة ولكس لامدا واختبار النسبة الفائية (F) في تحليل الدالة التمييزية أن جميع المتغيرات دالة إحصائياً ولها قدرة على التمييز، حيث كان المتغير الأكثر أهمية هو متغير معدل الثانوية، يليه معدل التخصيص، يليه اختبار القدرات، يليه الاختبار التحصيلي. وأشارت النتائج إلى كفاءة وجودة التصنيف لنموذج التحليل التمييزي حيث بلغت نسبة لتصنيف الصحيح له (90%).

وهذه النتيجة تتفق مع ما توصلت إليه كل من دراسة كارديان وأُلموس (Kardiyen & Olmus, 2016) التي أظهرت أن التحليل التمييزي ذو أفضلية في تصنيف البيانات، وكذلك دراسة ليدي آن وجاستون (Lydi-Anne & Gaston, 2011) التي توصلت إلى أن أسلوب التحليل التمييزي كان أكثر دقة في تصنيف متعاطي الكحول وغير المتعاطين، ودراسة ماروكو وآخرون (Maroco et al., 2011) التي أكدت أيضاً أن أسلوب التحليل التمييزي يتمتع بدرجة عالية في تصنيف البيانات.

ويمكن تفسير هذه النتيجة إلى أن ما يميز أسلوب التحليل التمييزي بأنه يتم فيه تحديد ابسط طريقة للتمييز بين المجموعات والتحقق من الفرق داخل المجموعات وبين المجموعات، ويسعى إلى تكوين نموذج إحصائي يصور العلاقات المتبادلة بين المتغيرات المختلفة وذلك من خلال إيجاد تركيبة خطية لمجموعة من المتغيرات يطلق عليها متغيرات التمايز (الشمراني، ٢٠٠٨)، كما أنه يعتمد على الوصول إلى دالة التمايز التي تعمل على تعظيم الفروق بين متوسط المجموعات وتقليل التشابه في أخطاء التصنيف في الوقت ذاته، وذلك من خلال إيجاد تجميعات خطية لمجموعة من المتغيرات (Johnson & Wicher, 2007)، إضافة إلى ذلك فإن بيانات التحليل التمييزي لا تحتاج لأن تكون معيارية، فيكون لها وسط صفري وتباين يساوي واحد، حيث أن نتيجة التحليل التمييزي لا تتأثر بتعيير مفردات المتغيرات (أبو علام، ٢٠٠٩).

**ثالثاً: النتائج المتعلقة بسؤال الدراسة الثالث والذي ينص على:** " أي الأسلوبين المستخدمين: التحليل التمييزي، الانحدار اللوجستي هو الأكثر دقة وفاعلية في تصنيف الطلبة المتعثرين دراسياً؟"

أشارت نتائج دقة تحليل البيانات إلى تفوق نموذج تحليل الانحدار اللوجستي على نموذج التحليل التمييزي في تصنيف الطلبة المتعثرين دراسياً، وذلك وفق محكات دقة التصنيف ونسبة التصنيف الصحيحة

والمساحة تحت منحنى ROC، بينما كانت الأفضلية لنموذج التحليل التمييزي على نموذج الانحدار اللوجستي في محك حساسية التصنيف. وبشكل عام كانت الأفضلية لنموذج الانحدار اللوجستي.

واتفقت هذه النتيجة مع دراسة الرواشدة (٢٠٢٢) والتي أكدت أفضلية أسلوب تحليل الانحدار في دقة التصنيف مقارنة بالتحليل التمييزي في تصنيف الطلبة الموهوبين. ودراسة برزنجي (Barznji. 2018) التي بينت أن تحليل الانحدار اللوجستي لديه أفضل تنبؤ لعوامل الخطر لمرض السكري. وذلك مع دراسة حسينة (٢٠١٨) التي أكدت أن نموذج الانحدار اللوجستي تفوق على التحليل التمييزي في التنبؤ بالمؤسسات الفاشلة.ودراسة ميهالوفيتش (Mihalovic, 2016) التي أكدت على نموذج الانحدار اللوجستي تفوق في دقة التصنيف على أسلوب التحليل التمييزي. ودراسة شايان وآخرون (Shayan et al., 2015) التي كان فيها أداء نموذج الانحدار اللوجستي أفضل من التحليل التمييزي في جميع المواقف عند استخدام مؤشري التصنيف Q و B. واختلفت هذه النتيجة مع نتائج ما توصلت له دراسة كل من ماروكو وآخرون (Maroco et al., 2011) التي أكدت أيضاً أن أسلوب التحليل التمييزي يتمتع بدرجة عالية في تصنيف البيانات، وكذلك دراسة ليدي آن وجاستون (Lydi-Anne & Gaston, 2011) التي توصلت إلى أن أسلوب التحليل التمييزي كان أكثر دقة في تصنيف متعاطي الكحول وغير المتعاطين.

يُفسّر التفوق النسبي لأسلوب تحليل الانحدار اللوجستي على نظيره أسلوب التحليل التمييزي الخطي على الرغم من كونه أقل شروطاً قد يعزى إلى الفلسفة الخاصة به التي تعتمد على أسلوب استخدامه للاحتمالات في معالجاته الخاصة بالمتغيرات التي يتعامل معها، وهي نفس الفلسفة التي تستند اليها نظرية الاستجابة للمفردات والتي أثبتت جُل الأدبيات نجاحها في تحقيق القياس الموضوعي في مجال السلوك الإنساني.

وتفسير هذه النتيجة بأن تحليل الانحدار اللوجستي أقل شروط ولا يقدم أي افتراضات حول توزيع المتغيرات المستقلة، حيث يستخدم عندما لا تتطلب المتغيرات المستقلة افتراض التوزيع الطبيعي (Sharma, 1996)، حيث أن التحليل التمييزي حتى يتم الحصول على نتائج صادقة نسبياً يتطلب: استخدام عينات من حجم متوسط أو كبير، وأن تتوزع المتغيرات التابعة توزيعاً اعتدالياً، وأن تكون تباينات وتغايرات المتغيرات التابعة في المجتمع واحدة في جميع مستويات العامل، وأن يتم اختيار العينة عشوائياً، كما أن درجة أي فرد في العينة في أي متغير يجب أن تكون مستقلة عن جميع درجات أفراد العينة الآخرين، لهذا لا يجب الثقة في اختبار الدلالة للتحليل التمييزي إذا انتهك شرط الاستقلالية (أبو علام، ٢٠٠٩)، لذلك يعتبر تحليل الانحدار اللوجستي أكثر قوة ومرونة في حال انتهاك هذه الافتراضات، ولا يضع هذا النموذج أي شروط مسبقة على المتغيرات التفسيرية، فيمكن أن تكون كمية مستمرة، أو كمية متقطعة، كما يمكن أن تكون نوعية أو وصفية، أو خليط من المتغيرات الكمية والوصفية.

ويرى لي (Lea, 1997) أنه وإن كانت هناك عدة أساليب إحصائية طورت لتحليل البيانات ذات المتغيرات التابعة التصنيفية، مثل تحليل الدوال التمييزية، إلا أن تحليل الانحدار اللوجستي يتمتع بعدة مميزات تجعله ملائما للاستخدام، ويشير ويلكر (Walker, 1996) أيضاً إلى أن تحليل الانحدار اللوجستي هو اقل حساسية تجاه الانحرافات عن اعتدالية التوزيع لمتغيرات الدراسة وذلك مقارنة بأساليب إحصائية أخرى مثل التحليل التمييزي، كما أن الانحدار اللوجستي يستطيع ان يتجاوز العديد من الافتراضات الشديدة لانحدار المربعات الدنيا الاعتيادية، الأمر الذي يجعل تحليل الانحدار اللوجستي هو الأسلوب الأفضل في حال المتغير التابع الثنائي.

**التوصيات**

- استخدام أسلوب تحليل الانحدار اللوجستي في الدراسات التنبؤية خاصة عندما يكون المتغير التابع (المتنبأ به) تصنيفياً ثنائياً.
- إعادة النظر في نصيب كل معيار من معايير القبول المعتمدة في جامعة القصيم، مع إعطاء وزن أكبر لكل من معدل الثانوية العامة، ومعدل التخصيص، والاختبار التحصيلي.
- إجراء دراسات علمية حول موضوع التنبؤ بالتعثر الدراسي باستخدام عوامل أخرى مثل: المتغيرات الديموغرافية، بعض المقررات التي يدرسها الطالب خاصة المشتركة، وغيرها.

**المراجع:**

**المراجع العربية**

أبو علام، رجاء محمود. (2009). *التحليل الإحصائي للبيانات باستخدام برنامج SPSS*. الطبعة الثالثة، دار النشر للجامعات.

الأستاذ، محمود، وصبح، أيمن. (2010). التعثر الأكاديمي وأسبابه لدى طلبة جامعة الأقصى ودور تكنولوجيا المعلومات والاتصالات في معالجته. *مجلة الجامعة الإسلامية (سلسلة الدراسات الإنسانية)*، 18 (1)، 39-81.

البنوي، علاء. (2021). *المقارنة بين فاعلية التحليل التمييزي والانحدار اللوجستي في دقة تصنيف البيانات تبعاً لخصائص العينة والمتنبئات* [رسالة دكتوراه غير منشورة]. جامعة اليرموك إربد، الأردن.

حسينة، جواني. (2018). *دراسة مقارنة بين الانحدار اللوجستي والتحليل التمييزي للتنبؤ بفشل المؤسسات، دراسة عينة من المؤسسات الاقتصادية لولاية ام البواقي* [رسالة ماجستير غير منشورة]، كلية العلوم الاقتصادية والتجارية وعلوم التسيير، جامعة العربي بن مهيدي، ام البواقي، الجزائر.

الخراشي، نهى بنت عبد الرحمن. (2021). معايير القبول في الجامعات وقدرتها التنبؤية في التحصيل الدراسي للطلاب والطالبات: جامعة الملك سعود نموذجاً. *مجلة شؤون اجتماعية*، 38(150)، 159–188.

الرواشدة، سكينة. (2022). *استخدام التحليل التمييزي والانحدار اللوجستي للكشف عن العوامل التي تساهم في تصنيف الطلبة الموهوبين والمتفوقين- دراسة مقارنة* [رسالة دكتوراه غير منشورة]. جامعة اليرموك إربد، الأردن.

الزامل، محمد بن عبدالله. (2012). قدرة معايير القبول على التنبؤ بالتقدم الأكاديمي لطلاب السنة التحضيرية بجامعة الملك سعود. *مجلة رسالة الخليج العربي*، 33(126)، 157–214.

الشافعي، محمد. (2013). القدرة التنبؤية بالمشاهدات ثنائية التصنيف لأسلوبي تحليل الانحدار اللوغاريتمي والتحليل التمييزي: دراسة مقارنة. *مجلة كلية التربية*، 37(2)، 92–134.

الشمراني، محمد. (2008). *دراسة مقارنة بين التحليل التمييزي وتحليل التباين في تحليل البيانات متعددة المتغيرات* [رسالة دكتوراه غير منشورة]. كلية التربية، جامعة أم القرى، المملكة العربية السعودية.

الشمراني، محمد. (2017). مقارنة بين أسلوبي انحدار الترجيح اللوغاريتيمي الثنائي وتحليل الدالة التمييزية في تحديد المتغيرات التمييزية وقوة التنبؤ بالأداء الأكاديمي بجامعة أم القرى. *مجلة العلوم التربوية*، 25(1)، 249–276.

صقر، عبد العزيز. (2003). مشكلات الشباب الحالية والمستقبلية كما يراها طلاب جامعة طنطا. *مستقبل التربية العربية*، العدد 23، المكتب الجامعي بالإسكندرية.

طاقية، البيومي؛ العتابي، كريم؛ والمنجي، هشام. (2016). استخدام الانحدار اللوجستي والتحليل التمييزي لدراسة حالات الاصابة بمرض الاسهال لدى الأطفال في العراق: دراسة تطبيقية. *المجلة المصرية للدراسات التجارية*، 40(1)، 233–255.

العليان، لولوة. (2017). العوامل المؤثرة في رسوب بعض طالبات المرحلة المتوسطة في المدارس الحكومية في مدينة عنيزة. *مجلة جامعة الفيوم للعلوم التربوية والنفسية*، (7)7، 325 – 376.

النعيمي، عزالدين. (2018). العوامل المنبئة بتسرب الطلبة ورسوبهم في جامعة الإمام محمد بن سعود الإسلامية باستخدام أسلوب الانحدار اللوجستي. *مجلة جامعة النجاح للأبحاث – العلوم الإنسانية*، 32 (11)، 2043 – 2074.

النعيمي، عزالدين. (2020). فاعلية أسلوب التحليل التمييزي في التنبؤ بالتعثر الدراسي من خلال اختبار القدرات العامة والاختبار التحصيلي في بعض الجامعات السعودية. *مجلة العلوم التربوية*، 24(2)، 460 – 524.

**المراجع الأجنبية**

Adler, W.; Brenning, A.; Potapov, S.; & Schmid, M. (2011). Ensemble Classification of Paired Data. *Computational Statistics & Data Analysis*. 55(5), 1933-1941.

Alkarkhi, A. F., & Easa, A. M. (2008). Comparing discriminant analysis and logistic regression model as a statistical assessment tools of arsenic and heavy metal contents in cockles. *Journal of sustainable development.*1(2), 102–106.

Antonogeorgos, G., Panagiotakos, B., Priftis, K., & Tzonou, A. (2009). Logistic Regression and Linear Discriminant Analyses in Evaluating Factors Associated with Asthma Prevalence among 10-12 years-old Children: Divergence and Similarity of the Two Statistical Method, *International Journal of Pediatrics*. https://doi.org/10.1155/2009/952042

Barznji, N. (2018). Using Logistic Regression Analysis and linear Discriminant Analysis to identify the risk factors of Diabetes. *Salahaddin university*. 22(6), 248-268, from https://www.researchgate.net/publication/333210883

Divjak, B., & Oreski, D. (2009). *Prediction of Academic Performance Using Discriminant Analysis*. paper presented in The ITI conference Cavtat Croatia, 225-230.

Ebrahimzadeh, F.; Hajizadeh, E.; Vahabi, N.; Almasian, M.; & Bakhteyar, K. (2015). Prediction of Unwanted Pregnancies Using Logistic Regression, Probit Regression and Discriminant Analysis. *Medical Journal*, Islam Repub Iran, V(29): 264.

Ferreras, P., & Feijoo. (2009). Scanning laser polarimetry: logistic regression analysis for perimetric glaucoma diagnosis, *Eye* (2009) 23, 593-600.

Gilbert, E. S. (1969). The effects of unequal variance-covariance matrices on Fisher's linear discriminant function, *Biometrics* 25, 505-515.

Gregori, D.; Petrinco, M.; & Simona, B. (2011). Using Data Mining Techniques in Monitoring Diabetes Care. The Simpler the Better, *Journal of Medical Systems*, 35(2), 277-281.

Hamid, H. (2010). A new approach for classifying large number of mixed Variables, *International Journal of Mathematical*, Computational, Physical, Electrical and Computer Engineering 4(10), 1355-1360.

Henrik Madeson. (2010). *Introduction to general and generalized linear models*, Uk, Champan and Hall/CRC, from https://www.cs.cmu.edu/~schneide/tut5/node42.html

Hosmer, D., Lemeshow, S., & Sturdivant, R. (2013). *Applied Logistic Regression*. U.S.A. John Wiley & Sons.

Huberty, C., & Olejnik, S. (2006). *Applied MANOVA and Discriminate Analysis*. (2nd ed), John Wiley & Sons, Hoboken, NJ.

Johnson, R. A., & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*. (6th ed), Pearson, Prentice Hall, New Jersey.

Josef, M. (2009). *Logistic Regression Models.* California institute of technology.

Kardiyen, F., & Olmus, H. (2016). A comparison of Two Group Classification Approaches to Fat tailed and Skewed Data. *Communication in Statistics- Simulation and Computation*, 45, 17-32.

Kutner, M.; Neter, J.; Li, W.; & Nachtsheim, C. (2005). *Applied Linear Statistical Models*. (5th ed). Irwin, McGraw-Hill, New York, U.S.A.

Lea. S. (1997). *Multivariate Analysis II: manifest variables analysis. Topic 4: Logistic Regression and Discriminant Analysis. University of Exeter Department of Psychology*. from http://www.exeter.ac.uk/~SEGLea/multivar/diclogihtml.

Lei, P., & Koehly, L. (2003). Linear discriminate analysis versus logistic regression: A comparison of classification errors in the two-group case. *Journal of Experimental Education*, 72(1), 25-49.

Lydi-Anne, V., & Gaston, G. (2011). Psychosocial determinants of intention to abstain from drinking alcohol while pregnant among a

sample of women of childbearing age. *Addiction Research and Theory*, 19(2), 128-137.

Maroco, J.; Silva, D.; Rodrigues, A.; Guerreiro, M; Santana, L.; & Mendonca, A. (2011). Data Mining Methods in the Prediction of Dementia: A Real - Data Comparison of the Accuracy, Sensitivity and Specificity of Linear Discriminant Analysis, Logistic Regression, Neural Networks, Support vector machines, Classification Trees and Random forests. *BMC Research Notes*, 4, 299, from https://www.researchgate.net/publication/51578344

Mihalovic, M. (2016). Performance Comparison of Multiple Discriminant Analysis and Logit Models In Bankruptcy Prediction. *Economic and Sociology*, 9(4), 101-118.

Moawed, S. A., Osman, M. M. (2017). The Robustness of Binary Logistic Regression and Linear Discriminant Analysis for the Classification and Differentiation between Dairy Cows and Buffaloes, International *Journal of Statistics and Applications.* 7(6), 403-310.

Musa, A.; Ibrahim, M.; Hamad, H.; & Shaheen, S. (2019). divergence and similarity of the binary logistic regression and linear Discriminant Analysis Models in Evaluating Factors Associated with Bluetongue Virus in Cattle. *International Journal of Statistics and Applications*. 9(6),180-185.

Pohar, M.; Blas, M.; & Turk, S. (2004). Comparison of Logistic Regression and linear Discriminant Analysis: A simulation study. *Metodoloski*, 1(1), 143-161.

Poorhaji, M.; Shahriyari, M.; Asadi, M.; Bahrami, F.; Gouvrchinghaleh, H.; Mohammadzadeh, T.; Rezayi, Z.; Navidian, E.; & Shankayi, Z,. (2022). Binary Logistic Regression and Linear Discriminant Analyses in Evaluating Laboratory Factors Associated with COVID-19: A Comparison of Two Statistical Methods. *Mediterranean Journal of Infection, Microbes & Antimicrobials.* 11(20), 152-158.

Santos, F.; Guyomarch, P.; & Bruzek, J. (2014). Statistical Sex Determination from Craniometrics: Comparison of Linear Discriminant Analysis, Logistic Regression. and Support Vector Machines. *Forensic Science International*. 245(205), 1-8.

Sharma, Subhash. (1996). *Applied Multivariate Techniques*, New York Chichester Brisbane Toronto Singapore.

Shayan, Z.; Mezerji, N.M.G.; Shayan, L.; & Naseri, P. (2016). Predication of Depression in Cancer Patients with Different Classification Criteria, Linear Discriminant Analysis versus Logistic Regression. *Global Journal of Health Science*, 8(7), 41-46.

Tabachnich, B. & Fidell, L. (2013). *Using Multivariate Statistics*. (6th ed). Pearson Educational, U.S.A.

Timen, N.H. (2002). *Applied multivariate analysis*, (2nd ed), Texts in statistics.

Walker, J. (1996). *Methodology Application: Logistic Regression Using the CODES Data*. Developed for Department of Transportation. National Highway Traffic Safety Administration (NHTSA), Washington DC.

Wan tang,. & Hue He (2012). *Applied Categorical and Count Data Analysis*, UK, champan and Hall/CRC.

William R. Klacka. (2010). *Discriminant Analysis*, UK, champan and Hall/CRC.