# تحليل تاثير بعض المتغيرات في رسوب الطلبة في مادة الرياضيات باستخدام نموذج الانحدار اللوجستي

**د. حازم منصور كوركيس سامي جبار شبوط**

جامعة بغداد/ كلية التربية – ابن الهيثم

## 1.1 الملخص

يهدف هذه البحث الى تحليل اثر بعض المتغيرات التي يعتقد انه تؤثر بشكل مباشر على رسوب طلبة الصف السادس العلمي في مادة الرياضيات، اذ استخدم الباحث اسلوب الانحدار اللوجستي ثنائي الاستجابة، وتم اختيار عينة عشوائية مكونة من(200) طالباً وطالبة، وقد درست حالة رسوب الطلبة وذلك باختيار عشرة متغيرات هي الجنس،مستوى الذكاء،الميول،هل طريقة التدريس جيدة،هل محتوى المادة ملائم،التحصيل الدراسي للاب،التحصيل الدراسي للام،عدد ساعات الدراسة،هل يتلقى دروس خصوصية،نوع الاسئلة، يعتقد انها المسبب الرئيسي لهذه المشكلة، اذ استخدم الباحث البرنامج الاحصائي SPSS في تحليل بيانات البحث، واستنتج الباحث ان ثلاث من هذه المتغيرات (مستوى الذكاء،هل يتلقى دروس خصوصية،عدد ساعات الدراسة) لها اثر معنوي في التأثير على رسوب الطلبة .

## المبحث الاول

## 1.2 المقدمة

تعد مادة الرياضيات من اهم المناهج التي يدرسها الطالب في الصف السادس العلمي لما لها من اهمية في تنمية قدراته العقلية على التحليل والاستنباط وفهم المعلومات ليس في الرياضيات فحسب بل في بقية المواد العلمية التي يدرسها الطالب في المقرر الدراسي. وتعد مشكلة ازدياد نسبة الرسوب في هذه المادة من المشاكل المهمة التي تواجه الطلبة في هذه المرحلة الدراسية. لذا يجب معالجة هذه المشكلة من خلال التعرف على اسباب فشل الطلبة في هذه المادة والقضاء على المعوقات بغية الارتقاء بالمستوى التعليمي للطلبة، اجريت العديد من الدراسات في المجال الاجتماعي والطبي ولكن هناك دراسات قليلة ونادرة في اسباب رسوب الطلبة، لذلك عمدنا الى اجراء هذا البحث للوقوف على اهم مشكلة تواجه الطلبة في مرحلة السادس العلمي والمتمثلة بتدني مستوى التحصيل في مادة الرياضيات وازدياد نسب الرسوب في هذه المادةلايجاد الحلول لهذه المشكلة للمساعدة على رفع المستوى العلمي للطلبة.

## 1.3 هدف البحث

يهدف البحث الى تحليل تاثير بعض المتغيرات في رسوب الطلبة في مادة الرياضيات باستخدام نموذج الانحدار اللوجستي ومعرفة قيم المعاملات ومدى تاثير هذه المعاملات في بناء هذا النموذج.

## 1.4اهمية البحث

تتركز اهمية هذا البحث في استخدام الانحدار اللوجستي كاداة مهمة في دراسة اسباب رسوب الطلبة وكيفية ايجاد الحلول المناسبة لهذه المعوقات وكيفية التغلب على هذه المشاكل لغرض رفع المستوى التعليمي للطلبة.

## 1.5 مشكلة البحث

لمعرفة اثر العوامل التي ادت الى ارتفاع نسب رسوب الطلبة في مادة الرياضيات للصف السادس العلمي تم دراسة المتغيرات التي تسببت في ارتفاع نسب الرسوب في مادة الرياضيات من اجل الوقوف على هذه المشكلة والوصول الى الطرق المقترحة لحل هذه المشكلة.

## 1.6 فرضية البحث.

يعتمد البحث على اختبار فرضية العدم التالية:

$H_0 = \beta_0 = \beta_1 \dots\dots \beta_{10} = 0$

باستخدام طرق الاختبار الواردة في متن البحث لمعرفة اي من المتغيرات التوضيحية لها تاثير معنوي على متغير الاستجابة.

معنوية معاملات الانحدار تساوي صفر في المجتع الذي سحبت منه العينة.

$H_0 = \beta_k = 0$ اي ليس للمتغيرات المستقلة من تاثير معنوي على المتغير المعتمد

## 1.7 ادوات البحث الاحصائية

بالنظر لكون المتغيرات التوضيحية ثنائية ومتعددة والمتغير المعتمد هو متغير ثنائي الاستجابة، تم استخدام اسلوب الانحدار اللوجستي وباستخدام الحزمة الاحصائية SPSS تم الحصول على النتائج المطلوبة.

## 1.8 حدود البحث

**الحدود البشرية:** 200 طالب وطالبة من الصف السادس العلمي

**الحدود المكانية:** بعض مدارس تربية الرصافة الثالثة

**الحدود الزمانية** : العام الدراسي 2017

## المبحث الثاني: الجانب النظري

### 2.1 مفهوم البيانات الثنائية

هناك العديد من الدراسات الاجتماعية والطبية والتي تهتم بتحليل ظاهرة ما وايجاد العوامل التي ترتبط بها وتؤثر عليها، من اجل التنبؤ بحدوثها او عدم حدوثها في ظل توافر معلومات عن تلك العوامل ومنها:[4]

وفي مثل هذه الحالات يكون المتغير التابع ثنائيا اما يساوي صفراً لعدم وقوع الحدث او واحد عند وقوع الحدث.[2]

ان اسلوب الانحدار الخطي المتعدد لايمكن استخدامه في مثل هذه الحالات لاسباب عديدة:[4]

1.تباين الخطأ غير ثابت ولايتبع التوزيع الطبيعي.

2. عدم امكانية تفسير القيم المتنبأ بها بوصفها احتمالات، حيث لايمكن حصر هذه القيم بين0و1.

لذا يمكن الاستعانة باساليب احصائية اخرى ومنها نموذج الانحدار اللوجستي.

### 2.2 أنموذج الانحدار اللوجستي Logistic Regression Model

#### 2.2.1 مفهوم الانحدار اللوجستي

يصنف النموذج اللوجستي ضمن النماذج اللاخطية التي يمكن تحويلها الى نماذج خطية , النماذج بالنماذج الخطية ضمنيا (جوهريا) (Intrinsically Linear model)[7]. وتسمى هذه ان نموذج الانحدار اللوجستي يعرف على انه احد النماذج التي تكون العلاقة بين المتغير التابع (y) والمتغيرات التوضيحية غير خطية، وتاخذ غالباً دالة الاستجابة الشكل S. وتعرف رياضياًكالاتي:[9]

$$\pi_i = E(Y_i|X_i) = \frac{\exp[\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots\beta_k x_{ik}]}{1+\exp[\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots\beta_k x_{ik}]}; i = 1,2 \ldots\ldots m \quad \ldots\ldots(1)$$

وباستخدام التحويل المسمى logit transformation تتحول الدالة غير الخطية في معادلة (4) الى دالة اخرى خطية وكالاتي:

$$\mathrm{L} = \ln\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \beta_k x_{ik} \; ; i = 1,2 \ldots\ldots m \qquad \ldots\ldots\ldots\ldots (2)$$

حيث ان $\frac{\pi_i}{1-\pi_i}$ *تمثل ما يسمى بنسبة الرجحان( (Odds Ratio) و هي نسبة احتمال حدوث الحدث الى* عدم حدوث ذلك الحدث[12]

وان L تمثل التحويلة $\mathrm{logit}(\pi_i)$.

ان معامل الترجيح (odds)هوعبارة عن اسلوب للتعبير عن مدى احتمال حدوث شئ ما مقارنة باحتمال عدم حدوثه، وغالباً مايتم التعبير عنه كما في المعادلة:[11]

$$odds = \left(\frac{\pi_i}{1-\pi_i}\right) = \exp[\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \beta_k x_{ik}] \ \ldots\ldots(6)\ ,0 < odds < \infty$$

ومن الجدير بالذكر ان معامل الترجيح (odds) قد حل مشكلة الحدود العليا للاحتمال (π) ، بحيث اصبح معامل الترجيح ياخذ اي قيمة من الصفر وحتى مالانهاية بدلا من π التي ينحصر مداها بين الصفر والواحد.

وبعد اخذ اللوغاريتم الطبيعي لدالة الترجيح يتسع المدى ليشمل كل الاعداد الحقيقية، اي ان: $-\infty < logit\pi_i < \infty$ ،واماابالنسبة لما يخص تفسير معاملات (معالم) نموذج الانحدار اللوجستي فهو يختلف تماما عن تفسير المعالم في نموذج الانحدار الخطي ، اذ يمثل مقدار التغير الحاصل في لوغاريتم نسبة الرجحان لحدوث الحدث نتيجة لتغير وحدة واحدة من المتغير المستقل وذلك بثبوت المتغيرات المستقلة الاخرى[14]

اما الوظيفة الرئيسية لدالة اللوجت هي السماح بتطبيق الانحدالر الخطي عند تحليل العلاقات للبيانات ذات المتغيرات التابعة الثنائية[13]

وتتميز تحويلةlogit بالعديد من الخصائص الازمة لشكل الانحدار الخطي حيث انها تتميز بما يلي:

1. دالة logit تكون مستمرة خطية في معالمها ( $\beta_0, \beta_1, \ldots\ldots \beta_k$).
2. كل *قيمة من قيم الطرف الايمن في* دالة logit المحددة بالفترة ( $-\infty, \infty$ ) تكون مقابلة لقيمة واحدة من *قيم*($\pi_i$) *المحددة بالفترة*(0,1).

### 2.2.2 بناء انموذج الانحدار اللوجستي

يبنى هذا النموذج على فرض اساس هو ان المتغير التابع (y) الذي نهتم بدراسته هو متغير ثنائي الاستجابة ياخذ القيمة (1) عند حدوث الاستجابة باحتمال $\pi_i$ والقيمة (0) عندعدم حدوث الاستجابة باحتمال 1-$\pi_i$، لهذا يكون المتغير التابع يتبع توزيع برنولي (Bernoulli distribution) وتكون دالة الكثافة الاحتمالية بالصيغة:[1]،[3]

$$p(Y = y_i) = {\pi_i}^{y_i}(1-\pi_i)^{1-y_i} \ldots\ldots(3)\ ;\ y_i = 0,1$$

لذا فأن توقع المتغير التابع يمثل احتمال حدوث الاستجابة:

$$E(y_i) = p(y = 1) = \pi_i \ldots\ldots(4)$$

اما تباين المتغير التابع حسب توزيع برنولي:

$$Var(y_i) = \pi_i(1-\pi_i) \ldots\ldots\ldots(5)$$

### 2.2.3 فروض نموذج الانحدار اللوجستي

بخلاف نموذج الانحدار الخطي فان نموذج الانحدار اللوجستي لا يتطلب التوزيع الطبيعي للمتغير المعتمد او تجانس التباين.

ان اساس الانحدار اللوجستي مبني على نظرية الاحتمال.

ان فرضيات الانحدار اللوجستي تنحصر في كون دالة تحويل logit خطية وان المنحني اللوغاريتمي الناتج لا يتضمن قيم شاذة، كما يجب ان يكون المتغير المعتمد تصنيفي (category) بحيث تكون هذه الاصناف مفصولة عن بعضها البعض تماماً، بحيث تتبع كل حالة صنفاً واحداً فقط[7].

### 2.2.4 تقدير معالم نموذج الانحداراللوجستي

في الانحدار الخطي يتم تقدير المعلمات با ستخدام طريقة المربعات الصغرى(O LS) والتي تهدف الى تصفير مربعات الخطا الى اقل مايمكن، اما في حالة الانحدار اللوجستي فيتم تقدير المعلمات باستخدام طريقة الامكان الاعظم (MLE) [14]

وهي من الطرق الاكثر ملائمة لكافة النماذج الخطية وغيرالخطية،وهي طريقة تكرارية (lterative ) تعتمد على تكرار العمليات الحسابية عدة مرات، حتى يتم الوصول الى افضل تقدير للمعلمات والتي يمكن تفسير البيانات المشاهدة.[10]

والصيغة الرياضية لدالة الامكان في حالة البيانات الثنائية تعرف بالشكل التالي:[12]

$$l(B_i) = \prod_{i=1}^{n} \pi_i{}^{y_i} (1-\pi_i)^{1-y_i} \dots\dots\dots (7)$$

وبالتعويض عن $\pi_i$ كما في معادلة رقم (4) بما يساويها وباخذ اللوغاريتم الطبيعي للطرفين وبالتبسيط نحصل على:

$$ln\big(l(B_i)\big) = ln \prod_{i=1}^{n} \pi_i{}^{y_i} (1-\pi_i)^{1-y_i} = \sum ln\, \pi_i{}^{y_i} (1-\pi_i)^{1-y_i}$$

$$= \sum y_i\, ln\pi_i + \sum (1-y_i)ln(1-\pi_i) = \sum ln(1-\pi_i) + \sum y_i\, ln \frac{\pi_i}{1-\pi_i}$$

$$= \sum y_i\,(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \beta_k x_{ik}\,)$$

$$- \sum \ln(1 + \exp[\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \beta_k x_{ik}\,]) \dots\dots (8)$$

وللحصول على المعلمات والتي هي تعظم دالة الامكان الاعظم(ML) نشتق الدالة الامكان بالنسبة للمعلمات التي يراد تقديرها ومن ثم جعلها تساوي صفراً فينتج (k+1) من المعادلات غير الخطية والتي لانستطيع حلها الا باستخدام خوارزمية تكرارية[6].

### 2.2.5 الاختبارات الاحصائية

في الانحدار اللوجستي هناك مجموعة من الاختبارات الاحصائية منها:

1. **اختبار *احصاءة* Wald** [5],[3]

تستخدم احصاءة Waldلبيان اهمية معاملات الانحدار اللوجستي ويعبر عن هذه الاحصاءة بصيغة المعادلة:

$$\text{Wald}=\left[\frac{\widehat{\boldsymbol{\beta}}}{\boldsymbol{S.E}(\widehat{\boldsymbol{\beta}})}\right]^{\mathbf{2}} \ldots\ldots\ldots\ldots(9)$$

حيث ان:

$\widehat{\boldsymbol{\beta}}$**:**تمثل **(**قيمة معامل الانحدار اللوجستي للمتغير المستقل X**)**

$\boldsymbol{S.E}(\widehat{\boldsymbol{\beta}})$**:** تمثل **(**قيمة الخطا المعياري لمعامل الانحدار اللوجستي للمتغير المستقل X**)**

ولها توزيع مربع كاي $\chi^2$ ،حيث تقارن القيمة الاحتمالية لاحصاءة( Wald(Significance مع مستوى المعنوية والتي تحدد مسبقا من قبل الباحث،لكي يتم التعرف فيما اذا كان المتغير معنوي ام لا ، و يكون هذا المتغير معنويا اذا كانت القيمة الاحتمالية لاحصاءة (Significance) Wald اقل من مستوى المعنوية والتي حددت مسبقاً ، ومن الجدير بالذكر ان احصاءة Wald يمكن ان تعاني قصورا شديدا اذا كانت القيمة المطلقة لمعامل الانحدار كبيرة ، الامر الذي ينتج عنه قيمة الخطا المعياري ستكون كبيرة جدا مما ينتج عنه قيمة صغيرة لهذه الاحصاءة ، وهذا بدوره سيؤدي الى جعل نتيجةهذا الاختبار للمتغير المعني ليست معنوية.

2. **اختبار Hosmer And Lemeshow** [8]

يستخدم هذاالاختبار اذاكان النموذج المستخدم يمكن ان يمثل بيانات الدراسة *بشكل* جيد *ام لا، اذ يستخدم اختبار مربع كاي* $\chi^2$ *لحسن المطابقة لتقييم الفرق بين القيم المشاهدة* ( observed)*وبين القيم المتوقعة* (Expected) واختبار الفروض التالية:

$H_0$ :*تساوي الحالات المشا*هدة مع الحالات المتنبا بها ، اي ان النموذج يمثل البيانات بشكل جيد.

$H_1$ *: عدم تساوي الحالات المشاهدة مع الحالات المتنبا بها ، اي ان النموذج لا يمثل البيانات بشكل* جيد.

فاذا كانت القيمة الاحتمالية لاحصاءة لمربع كاي (*Significance*) اكبر من مستوى المعنوية التي حددهاالباحث فيمكن القرار بقبول فرضية العدم $H_0$.

## 3. التطبيق العملي

### الرموز المستخدمة لمتغيرات البحث

في هذه الدراسة اقترحنا ان:

$Y$: المتغير المعتمد(متغير الاستجابة) : وهو متغير نوعي ثنائي ياخذ احدى القيمتين ،*(1)* اذا كان الطالب راسبا(فاشلا) في مادة الرياضيات،*(0)* اذا كان الطالب ناجحا.

اما المتغيرات التوضيحية فهي تمثل العوامل المؤثرة على المستوى العلمي للطلبة في مادة الرياضيات وهي كالاتي:

1. الجنس (*0* الانثى، *1* الذكر) $X_1$
2. مستوى الذكاء (*0* ضعيف، *1* متوسط، *2*عالي) $X_2$
3. الميول (*0* لا يوجد،*1* يوجد) $X_3$
4. هل طريقة التدريس جيدة (*0* كلا، *1*نعم) $X_4$
5. هل محتوى المادة ملائم (*0*كلا ، *1*نعم ) $X_5$
6. التحصيل الدراسي للاب (*0*امي،*1*ابتدائي،2متوسط،3اعدادي،4جامعي فما فوق ) $X_6$
7. التحصيل الدراسي للام (0امي،1ابتدائي،2متوسط،3اعدادي،4جامعي فما فوق) $X_7$
8. عدد ساعات الدراسة (0 قليلة، 1 متوسطة، 2 كثيرة ) $X_8$
9. هل يتلقى دروس خصوصية (0 كلا ، 1 نعم ) $X_9$
10. نوع الاسئلة (0 سهلة، 1 متوسطة، 2 صعبة ) $X_{10}$

### 4.الدراسة التطبيقية

باستخدام الحزمة البرمجية الجاهزة *SPSS* اوباستخدام ايعاز *Enter* وبعد ترميز وتجهيز البيانات وادخالها الى الحاسوب وثم اختيار الامر *Analyze* ثم الامر الفرعي *Regression* ثم الامر الفرعي *Binary logistic* ثم الامر *OK* حصلنا على المعلومات الوصفية لعينة الدراسة الملخصة في جدول رقم. *(1)*

**Case Processing Summary**

جدول رقم (1)

| | | N | Percent |
|---|---|---|---|
| Selected Cases | Included in Analysis | 200 | 100.0 |
| | Missing Cases | 0 | .0 |
| | Total | 200 | 100.0 |
| Unselected Cases | | 0 | .0 |
| Total | | 200 | 100.0 |

يلخص جدول رقم (1) البيانات التي تم ادخالها افي التحليل وكذلك حجم العينةالتي قمنا بدراستها و(*Missing data* ) البيانات المفقودة .

**جدول رقم (2)**

**Dependent Variable Encoding**

| Original Value | Internal Value |
|---|---|
| نجاح | 0 |
| فشل | 1 |

يمثل الجدول رقم (2) رموز او (*cods*) قيم المتغير التابع.

**جدول رقم (3)**

**Iteration History**

| Iteration | | -2 Log likelihood | Coefficients | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Constant | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ | $X_9$ | $X_{10}$ |
| Step 1 | 1 | 178.140 | 3.519 | 0.443 | -2.036 | -0.158 | -0.682 | 0.271 | 0.008 | -0.048 | -0.577 | -0.732 | -0.187 |
| | 2 | 159.130 | 5.436 | 0.601 | -3.214 | -.0182 | -0.887 | 0.274 | -0.003 | -0.086 | -0.964 | -1.121 | -0.157 |
| | 3 | 154.955 | 6.833 | 0.675 | -4.179 | -0.181 | -0.941 | 0.217 | -0.020 | -0.095 | -1.198 | -1.341 | -0.143 |
| | 4 | 154.327 | 7.531 | 0.696 | -4.752 | -0.172 | -0,941 | 0.188 | -0.028 | -0.095 | -1.265 | -1.405 | -0.145 |
| | 5 | 154.290 | 7.721 | 0.698 | -4.931 | -0.170 | -0.940 | 0.184 | -0.029 | -0.095 | -1.271 | -1.411 | -0.146 |
| | 6 | 154.290 | 7.737 | 0.698 | -4.946 | -0.170 | -0.940 | 0.184 | -0.029 | -0.095 | -1.271 | -1.411 | -0.146 |
| | 7 | 154.290 | 7.737 | 0.698 | -4.946 | -0.170 | -0.940 | 0.184 | -0.029 | -0.095 | -1.271 | -1.411 | -0.146 |

يتضمن جدول رقم (3) عدد الدورات التكرارية لمشتقات دالة الامكان الاعظم (ML) والتي من خلالها يتم الحصول على اقل قيمة لسالب ضعف لوغاريتم دالة الامكان الاعظم ( 2- Log likelihood) والذي يمثل التقدير الامثل لمعالم النموذج.

ونلاحظ من الجدول اننا قد حصلنا في الدورة السابعة لمشتق سالب ضعف دالة الامكان الاعظم (-2 Log likelihood) على اقل قيمة (154.290)، ,وتم التوقف عند هذه الدورة وذلك لان التغير في المعاملات $(P_1, P_2, \dots\dots\dots, P_{10})$ اصبح اقل من0.001 ، وفي حقيقة توقفنا عند هذه الدورة لان التغير في المعاملات بعد الدورة الرابعة اصبح بطيئاجداً،ونلاحظ من الجدول بان مقدرات المعالم في الدورات *5,6,7* هي متشابهة مع فروقات قليلة جدا واعتبرنا معالم الدورة السابعة افضل نتيجة يمكن الحصول عليها وتوقفنا عندها، اذ ان ( 2-) Log likelihood عند هذه الدورة في نهايتها الصغرى.

Variables in the Equation

**جدول رقم (4)**

| | B | S.E. | Wald | Df | Sig. | Exp(B) | 95.0% C.I.for EXP(B) | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Lower | Upper |
| Step1 X1 | 0.698 | 0.422 | 2.729 | 1 | 0.099 | 2.009 | 0.878 | 4.599 |
| X2 | -4.946 | 1.078 | 21.058 | 1 | 0.000 | 0.007 | 0.001 | 0.059 |
| X3 | -0.170 | 0.471 | 0.130 | 1 | 0.719 | 0.844 | 0.335 | 2.124 |
| X4 | -0.940 | 0.707 | 1.767 | 1 | 0.184 | 0.391 | 0.098 | 1.562 |
| X5 | -0.184 | 0.507 | 0.131 | 1 | 0.717 | 1.202 | 0.445 | 3.247 |
| X6 | -0.029 | 0.164 | 0.031 | 1 | 0.859 | 0.971 | 0.704 | 1.341 |
| X7 | -0.095 | 0.182 | 0.273 | 1 | 0.601 | 0.909 | 0.637 | 1.298 |
| X8 | -1.271 | 0.385 | 10.883 | 1 | 0.001 | 0.280 | 0.132 | 0.597 |
| X9 | -1.411 | 0.448 | 9.931 | 1 | 0.002 | 0.244 | 0.101 | 0.587 |
| X10 | -0.146 | 0.409 | 0.128 | 1 | 0.721 | 0.864 | 0.388 | 1.925 |
| Constant | 7.737 | 1.913 | 16.363 | 1 | 0.000 | 2292.458 | | |

يلخص جدول رقم (4) معالم النموذج الامثل والتي *حصلنا عليها* في الدورة السابعة (من الجدول رقم (3))

*ويتضمن جدول رقم(4) ،جميع معالم النموذج المقدرة* $(\beta_0, \beta_1, \ldots\ldots \beta_{10})$ وقيمة الخطا (S.E)المعياري لكل معلمة واحصاءة *Wald* لكل معلمة من معلمات النموذج وكذلك عدد *درجات* الحرية ومعنوية المعلمات(Sig) والتي يفسر الجدول على النحو التالي:

ان العمود ( B) يحتوي على معاملات النموذج والتي هي بدلالة Log-odds ويمكن كتابة النموذج والتي كانت معاملات الانحدارللمتغيرات المستقلة ذات تاثير على المتغير المعتمد على النحو التالي $\text{Log}\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = 7.737 - 4.946x_2 - 1.271\ x_8 - 1.411x_9$ :

**اذ ان :**

$\hat{\pi}$ هي تمثل احتمال الحصول على نتيجة (فشل) للقرارات الجديدة والتقديرات هذه توضح العلاقة بين المتغيرات التوضيحية والمتغير المعتمد بوحدات logit.

و العمود الثاني فهو يمثل عمودالخطا المعياري للمعاملات (S.E) والذي يكون وفق العلاقة :

$$\text{S.E}(\hat{B}_i) = hii$$

اذان $hii$ تمثل العناصر القطرية (Diag) لمصفوفة التباين المشترك المقدرة وفق العلاقة :

$$Cov(\hat{B}_i) = [x' Diag(n_i, \hat{\pi}_i\ (1-\hat{\pi}_i)x]^{-1}$$

و العمود الثالث فيمثل احصاءة Wald لاختبار معنوية المعاملات (كما في معادلة رقم(9))

و يتوزع وفقاً لتوزيع $\chi^2$ بدرجة حرية df=1

اما بالنسبة لفرضية العدم(الفرضية الصفرية) لكل معامل من المعاملات فهي $H_0 : \beta_i = 0$

بينما تكون الفرضية البديلة $H_1 : \beta_i \neq 0$، وهو اختبار من طرفين.

اما بالنسبة Sig فهو عمود يمثل معنوية المعاملات المناظرة لقبول او رفض فرضية العدم وذلك باستخدام الاحتمالات ، فاذا كانت Sig<0.05 (.05 =0 α ) فان المعامل يكون معنوي وهو لايساوي صفر في المجتمع الذي سحبت منه هذه العينة.

اما العمود (Exp(B) odds Ratio) فانه يوضح قيمة الدالة الاسية لمعامل الانحدار وهو يعبر عن المضاعف الذي تتغير به نسبة الترجيح .

اما بالنسبة الى C.I for Exp(B) فيمثل عمود حدود الثقة.

لا اختبار كفاية النموذج بالكامل وجودته ( *Goodness of fit*) في حالة الانحدار الخطي كنا نستخدم احصاءة $F$ واحصاءة$R^2$ ، اما في حالة الانحدار اللوجستي فنستخدام نسبة الامكان الاعظم *log likelihood Ratio*والذي يتبع توزيع مربع كاي( $\chi^2$ ) وفقاً للعلاقة:[2],[4]

$$\chi^2 = 2[\log L_0 - \log L_1]$$

*حيث ان:*

$L_1$ : *قيمة دالة الامكان الاعظم التي تحتوي* على (i) من المتغيرات.

$L_0$ : *قيمة دالة الامكان الاعظم التي تحتوي* على (i−1)من المتغيرات .

نلاحظ من الجدول رقم(5) ان قيمة $\chi^2$ =122.969 *وهي معنوية عند مستوى دلالة* ($\alpha < 0.001$) اذ ان Sig=0.000،وبدرجة حرية df=10 (عدد المتغيرات التوضيحية) مما يؤكد ذلك معنوية النموذج الموفق بالكامل .

**جدول رقم (5)**

**Omnibus Tests of Model Coefficients**

| | Chi-square | df | Sig. |
|---|---|---|---|
| Model | 122.969 | 10 | .000 |

**جدول رقم(6)**

**Contingency Table for Hosmer and Lemeshow Test**

| | | group =النجاح | | group = الفشل | | |
|---|---|---|---|---|---|---|
| | | Observed | Expected | Observed | Expected | Total |
| Step 1 | 1 | 20 | 19.921 | 0 | 0.079 | 20 |
| | 2 | 18 | 18.615 | 2 | 1.385 | 20 |
| | 3 | 15 | 15.725 | 5 | 4.275 | 20 |
| | 4 | 14 | 14.312 | 6 | 5.688 | 20 |
| | 5 | 15 | 11.706 | 5 | 8.294 | 20 |
| | 6 | 10 | 9.972 | 11 | 11.028 | 21 |
| | 7 | 6 | 6.859 | 14 | 13.141 | 20 |
| | 8 | 2 | 2.500 | 18 | 17.500 | 20 |
| | 9 | 0 | 0.325 | 20 | 19.675 | 20 |
| | 10 | 0 | 0.065 | 19 | 18.935 | 19 |

يمثل جدول رقم (6) ايضاًاختباراً لا معلمياً لجودةالتوفيق للنموذج ،اذ يعتمد هذا الاختبارعلى حساب احصاءة مربع كاي$X^2$ للفرق بين القيم المشاهدة( observed) والقيم المتوقعة ( Expected) وقد تم اقتراحها من قبل كل من ( Hosmer and Lemeshow) للكشف عن انحرافات النموذجاللوجستي وذلك باستخدام توزيع مربع كاي ($X^2$ distribution)، وتتكون احصاءة الاختبارهذا من جزئين، جزء مشاهد observed لايستند الى نموذج نظري والجزءالاخر متوقع Expected يكون محسوب من تقديرات النموذج اللوجستي.

ويتم احتساب احصاءة مربع كاي$\chi^2$ لجودة التوفيق من تقاطع المجاميع للمتغير المعتمد الثنائي (Y) مع مجاميع من الاحتمالات التقديرية،، وتستخدم احصاءة (H ) التي لها توزيع مربع كاي $X^2$ ،وذلك لاختبار معنوية الفروقات بين التكرارات الفعلية( observed) وبين التكرارات المتوقعة( Expected)، عندما يكون اسلوب التجزئة محددا بنقاط ثابتة ضمن المدى [0,1] ،حيث يمكن اختيار اي عدد من نقاط هذه التجزئة، وغالبا ما تكون نقاط التجزئة: (m=10 )وفي هذه الحالة تكون المجموعة المتضمنة ادلة الازواج( $Y_i P(X_i)$) ضمن المجموعة k وفق الاتي:[8],[7]

$$Jk=\{i:(k-1\backslash m\leq P(X_i) \leq k\backslash m\}$$

*وتحسب التكرارات المشاهدة والمتوقعة في صفي العمود k كالاتي:*

$$\hat{h}_{1k} = \sum_{i\in jk} Yi = \sum_{i\in jk} \hat{P}\,(X_i)$$

$$\hat{h}_{2k} = \sum_{i\in jk} [\ 1 - Yi] = \sum_{i\in jk} [\ 1 - \hat{P}\,(X_i\,)]$$

*وبالنسبة الى فرضية العدم*

$$H_0 = h_{1k} = \hat{h}_{1k}$$
$$= h_{2k} = \hat{h}_{2k}$$

*وان احصاءة اختبار H تحسب وفق* العلاقة[8],[7]:

$$H=\sum_{s=1}^{2}\sum_{j=1}^{m}(h_{sj} - \hat{h}_{sj})/\hat{h}_{sj}$$

وان احصاءة H تتوزع وفق توزيع مربع كاي ($X^2$ distribution) وبدرجة حرية ( df =m-2).

كما نلاحظ ذلك من جدول رقم (7):

**جدول رقم (7)**

**Hosmer and Lemeshow Test**

| Step | Chi-square | df | Sig. |
|---|---|---|---|
| 1 | 3.463 | 8 | .902 |

اذ ان قيمة احصاءة مربع كاي كانت (3.463 )، ومن ثم نقبل فرضية العدم لان (Sig=0.902) من اجل درجة الحرية df=8، وهذا ما يؤكد جودة التوفيق للنموذج بالكامل والذي يؤكده الجدول رقم (6) للعمود k ،وذلك للقيمتين (1, 0) (راسب ,ناجح) بين القيم الفعلية والقيم التقديرية.

**جدول رقم(8)**

**Classification Table**

| Observed | | | Predicted | | |
|---|---|---|---|---|---|
| | | | Group | | |
| | | | النجاح | الفشل | Percentage Correct |
| Step 1 | group | النجاح | 84 | 16 | 84.0 |
| | | الفشل | 18 | 82 | 82.0 |
| | | Overall Percentage | | | 83.0 |

يمثل الجدول رقم (8) جدول التصنيف (Classification Table)، حيث يبين هذا الجدول النسبة المئوية للتصنيف الصحيح overall percentage = 83% الى مجموعتي التصنيف التي تنتمي اليها و التي وجدت بالصورة 0.83= 200/(82+84) وهي نسبة جيدة والتي تدل على ان النموذج يمثل البيانات بشكل جيد.

## 5.الاستنتاجات

من الجدول رقم (4) نلاحظ مايلي:

**1**-احتل المتغير $X_2$(مستوى الذكاء) المرتبة الاولى بالتاثير على المتغير المعتمد (Y) (رسوب الطلبة ) اذ ان معامل الانحدار لهذا المتغير $\beta_2 = -4.946$ وان هذا المعامل اظهر معنوية عالية sig=0.000 على المتغير التابع عند مستوى دلالة ($\alpha > 0.001$ ) بدرجة حرية (df=1) وان احصاءة Wald=21.058.

2- في المركز الثاني احتل المتغير$X_9$ (والتي تلقي الطالب الدروس الخصوصية) من حيث الاهمية في التاثير على المتغير(Y) اذ ان معامل الانحدار لهذا المتغير $\beta_9 = -1.411$ وهو يفسربان المتغيرالتوضيحي سينخفض وحدة واحدة من احتمال رسوب الطالب بمقدار (411. 1) وهو معنوي sig= 0.002 وان احصاءة931. Wald=9 .

3- احتل المتغير ($X_8$عدد ساعات الدراسة) المركز الثالث في التاثير على المتغير المعتمد اذ كان معامل الانحدار لهذا المتغير $\beta_9$ = −1.271 وهو معنوي sig= 0.001 وان احصاءة كانت Wald=10.883.

4- اما المتغيرات.$X_1$(االجنس )،$X_3$(االميول )،$X_4$( طريقةالتدريس)،$X_5$( محتوى االمادة )، (المستوى التعليمي للاب)$X_6$ ،(المستوى التعليمي للام) $X_7$ ( نوع الاسئلة) $X_{10}$ فكانت غير معنوية في التاثير على المتغير المعتمد.

وفي هذا الصدد نود الاشارة الى مايلي:

1- اظهرت نتائج البحث تدني المستوى العلمي لطلاب الدراسة الثانوية في مادة الرياضيات وزيادة عددهم لذلك يجب العمل على رفع المستوى العلمي للطلبة من خلال متابعة الاهالي لابنائهم ، وكذلك مدرسين المادة من خلال توعية الطلبة لكي يكون للطالب ميول وحب لهذه المادة.

2- اجراء دراسات مستمرة والتعرف على المشاكل التي يمر بها الطلبة من خلال كثرة الزيارات الميدانية للمدارس والوقوف عن كثب للتعرف على المعوقات التي تواجه الطلبة.

3- توسيع استخدام الانحدار اللوجستي ثنائي الاستجابة في العديد من المجالات التربوية والطبية والاجتماعية.

## 6.التوصيات

من خلال هذا البحث قد تم تسليط الضوء على تحليل الانحدار اللوجستي ثنائي الاستجابة وقد تمحور على اتجاهات وهي:

1- الدراسة النظرية لمفهوم الانحدار اللوجستي وامكان تقدير معالمه باستخدام طريقة الامكان الاعظم.

2- تحديد الاسباب والعوامل التي تؤثرفي رسوب الطلبة باستخدام تقنية الانحدار اللوجستي.

3- ان استخدام تقنية الانحدار اللوجستي ساعد في تحديد المتغيرات التي تزيد او تخفض من اسباب فشل او نجاح الطلبة.

## 7.المقترحات

1- يمكن لاستفادة من نموذج الانحدار اللوجستي في نمذجة المتغيرات االتابعة التي تكون ثنائية القيمة لما تمتاز به هذه المتغيرات من قوة تفسيرية كبيرة.

2- ضرورة توسيع استخدام الانحدار اللوجستي في كافة المجالت الطبية والاجتماعية والاقتصادية .

3-اجراء دراسات تطبيقية لنماذج الانحدار اللوجستي للمتغيرات ذات المستويات الرتبية والمتعددة.

## 8 .المراجع

1. البلداوي، تسنيم حسن 1996م"مقارنة تحليلية بين انموذج الانحدار اللوجستي ونماذج الدوال التمييزية" اطروحة دكتوراه (غير منشورة)،كلية الادارة والاقتصاد،جامعة بغداد.

2. الجاعوني ،فريد وعدنان غانم (2011) "استخدام تقنية الانحدار اللوجستي ثنائي الاستجابة في دراسة اهم المحددات الاقتصادية والاجتماعية لكفاية دخل الاسرة،دراسة تطبيقية على عينة عشوائية من الاسر في محافظة دمشق" بحث منشور في مجلة جامعة دمشق للعلوم الاقتصادية والقانونية، المجلد7،العدد الاول.

3. الراوي خاشع محمود(1978)"المدخل الى تحليل الانحدار"كلية الزراعة والغابات،جامعة الموصل.

4.الدليمي،.محمد ناجي و.اموري هادي كاظم(1990)"تحليل الانحدار بالامثلة"جامعة بغداد.

5. فهمي،محمد شامل(2005) " الاحصاء بلا معاناة،المفاهيم مع التطبيقات باستخدام برنامج "SPSS" الجزء الثاني،مركز البحوث، المملكة العربية السعودية.

6.D.Cook,P.and others(2001), "Binary Response and Logistic Regression Analysis" part of the lowa state University NSE/ILI project Traditional Statistical Methods,www.public .iastate.edu.

7.Draper,N.R.and Smith, ,H.(1981)"Applied Regression Analysi s" New York,p.413.

8.Hosmer ,D.W.,and Lemeshow,H.(1988)"Goodnees of fit testing for the Logistic model when the estimated probabilities are small"

9.Kleinbaue,D.G Kupper,L.Land Muller.K.E(1988):Applied Regression Analysis and other multibariable methods DWS-KENT publishing company,a division of wadsworth P317.

10.Newsom,R.(2003)"Data AnalyisisII: Logistic Regression"Fall 2003.

11.Pample,C.(2000) " *Logistic Regression Aprimer*" sage university paper.

12.Ret,K and David A. (1983) "Non Linear Regression Modeling" Marcel Dekker,NewYork.

13.Wallker,D.(1996) "Discriminant Funcation Analysis" lesson 8.

14.(2010) "Logistic Regression " ,http//pcs.maths.lancs.ac.uk.

## Analysis of the effect of some variables in the students failure in mathematics using logistic regression model

### Abstract

The aim of this study is to analysis the effect of some variables that are believed to have a direct effect on the failure of students of the sixth grade scientific in mathematics subject. The researcher used the method of logistic regression with binary response. A random sample of 200 students was selected. Ten variables are, namely, The sex, The level of intelligence, the tendencies, is the teaching method good?, is the content of the subject appropriate?, the educational achievement of the father, the academic achievement of the mother, the number of hours of study, receive private lessons, type of questions. The researcher used the statistical program SPSS in the analysis of the studied data. The researcher concluded that three of these variables( The level of intelligence, does he receive tutoring, the number of hours of study ) have a significant effect on the student failure.