تعيين بعض المتغيرات المؤثرة على منح شهادة أساسيات الحاسوب والانترنت (IC3)

مروان عبد العزيز دبدوب **أحمد محمد زكي إسماعيل**

كلية علوم الحاسوب والرياضيات

جامعة الموصل

تاريخ استلام البحث: ٢٠١٢/٠٩/١٧ **تاريخ قبول البحث: ٢٠١٣/٠١/٣٠**

ABSTRACT

To identify certain variable that influence the result Internet and Computing Core Certification examination (IC3) which is conducted in the Computer Center and Internet at the University of Mosul. The affect of the response of the examination score in the computer applications by seven explanatory variables were studied, the variables are: the score of computer basics examination, the examinee's sex, the experience of the examinee in the computer, the specialization of the examinee, the participation of the examinee in courses, degree to be obtained through study and the repeated participation in the examination.

Four statistical procedures were applied to select the best regression equation that involves the explanatory variables which have a real effect on the dependent variable. These procedures include: forward selection, backward elimination, stepwise regression and all possible procedures depending on certain comparison measures, which include mean square error MSe, and Mallows-Cp Statistics.

It was clear that the best model which non-presence of the constant β_0, and involves the variables: the score of computer basics examination, the examinee's sex, the specialization of the examinee and degree to be obtained through study.

الملخص

للتعرُّف على المتغيرات المؤثرة على درجة امتحان التطبيقات الأساسية وهو أحد الاختبارات للحصول على شهادة أساسيات الحاسوب والانترنت (IC3) والذي يُجرى في مركز الحاسوب والانترنت في جامعة الموصل، فقد دُرس تأثر درجة امتحان التطبيقات الأساسية (المتغير المعتمد) بسبعة متغيرات توضيحية هي: درجة الامتحان في أساسيات الحاسوب، جنس المُمتحن، خبرة المُمتحن بالحاسوب، تخصص المُمتحن وفق الدراسة الجامعية، مشاركة المُمتحن بالدورات، الدراسة المتقدم إليها وإعادة المشاركة بالامتحان.

طُبِّقت أربع طرائق إحصائية وقورن بينها وهي: طريقة الاختيار الأمامي، الحذف العكسي، الانحدار المتدرج وطريقة كل الانحدارات الممكنة، وذلك لاختيار أفضل أنموذج انحدار يضم المتغيرات المؤثرة فعلياً على المتغير المعتمد بالاعتماد على مقياسي المفاضلة: متوسط مربعات الخطأ وإحصائية مالو.

اتضح أن الأنموذج الأفضل هو الذي لا يحتوي على الثابت β_0 ويضم المتغيرات: درجة الامتحان في أساسيات الحاسوب، جنس المُمتحن، تخصص المُمتحن وفق الدراسة الجامعية، الدراسة المتقدم إليها.

المُقدّمـــــــة

منذ بضعة سنين ادخل منح شهادة أساسيات الحاسوب والانترنت Internet and Computing Core Certification (IC3) كأحد متطلبات الانخراط في احد مستويات الدراسات العليا لنيل شهادة: دبلوم عـــالي، ماجستير أو دكتوراه. يتـــضمن امتحـــان IC3 ثلاثـــة اختبـــارات هـــي: مبـــادئ الحاســـوب Computing Fundamental والتطبيقـــــات الأســـاســـية Key applications والتعامل مـــع الانترنيـــــت Living on line، وهنالك الكثير من المراكز والمعاهد والشركات أبدت اهتماماً بهذا الموضوع مثل شركة الخليج للتدريب والتعليم (2005) [5]. توجد عوامل مؤثرة على كل أو أحد اختبارات IC3، فتحليل الانحدار مـــن التحاليـــل الإحصائية المعتمدة على منهجية اختبار الفرضيات لاختيار العوامل المؤثرة وتوضيح العلاقة بينها وبين نتـــائج اختبارات IC3، فضلا عن التنبؤ والسيطرة.

إن تحليل الانحدار يُعطي طريقة سهلة مفهومة لبحث العلاقات الدالية بين المتغيرات، ويـــستخدم فـــي وصف البيانات والسيطرة عليها وكذلك تقدير المعلمات والتنبؤ [9]. من الأفضل أن نختار أنموذج الانحدار الذي يحتوي على أقل عدد ممكن من المتغيرات التوضيحية ذات الأهمية الإحصائية في تفسير التغيرات في المتغير المعتمد وأن النجاح أو الفشل للأنموذج المُقدر يعتمد على اختيار طريقة تقدير معلمات الأنموذج ومن ثم علـــى صفات المعلمات التي سيتم تقديرها حسب متغيرات الدراسة[1].

الهـــدف

هنالك بعض المتغيرات التي يعتقد بأنها تؤثر على درجة امتحان التطبيقـــــات الأســـاســـية وهو أحد الاختبارات للحصول على شهادة أساسيات الحاسوب والانترنت (IC3)، ولتعيين المتغيرات المؤثرة ستطبق أربع طرائق إحصائية في تحليل الانحدار وهي: طريقة الاختيار الأمامي، الحذف العكسي، الانحدار المتدرج وطريقة كل الانحدارات الممكنة. ومن خلال المقارنة بين نتائج الطرائق الأربع سيستخلص أفضل أنموذج انحدار يضم المتغيرات المؤثرة فعلياً على درجة امتحان التطبيقـــــات الأســـاســـية بالاعتماد على مقياسي المفاضلة: متوسط مربعات الخطأ (mean squares error-MSe) وإحصائية مالو (Cp-Mallow's statistics).

تحليل الانحدار

إن تحليل الانحدار يتمثل بعلاقة سببية بين متغير أو مجموعة متغيرات يُطلـــق عليهـــا بـــالمتغيرات التوضيحية أو المُفسرة (Explanatory Variables) ومتغير معتمد (Dependent Variable).

إن الهدف من تطبيق تحليل الانحدار هو تحديد العوامل التي تؤثر في المتغير المعتمد من خلال دراسة مجموعة من البيانات التي تُساعد في تقدير معلمات الأنموذج ومن ثم التنبؤ أو تقدير قيمة المتغير المعتمد عنـــد توفر القيم التقديرية للمتغيرات التوضيحية، فإنَّ تحليل الانحدار يعطي قيماً عددية لتأثير العوامل المُفسرة علـــى المتغير المعتمد ويلاحظ من ذلك أن أنموذج الانحدار يعتمد دائماً على العلاقة السببية بمعنى أن يكون التغير في المتغير التوضيحي سبباً رئيسياً للتغير في المتغير المعتمد [10].

لقد ذكر الجاعوني (2008) [1] أنه من الضروري التفريق بين نوعين من العلاقات التي تربط المتغير المعتمد بالمتغيـــرات التوضيحيـــة، النوع الأول يسمى بالعلاقة الدالية $f(x)=y$ هذا يعني أن تغير قيم المتغيـــر المعتمد يعتمد فقط على تباين المتغيرات التوضيحية في الأنموذج، والنوع الثاني يسمى بالعلاقة الإحصائية التي

أنموذجها لا يمكنه احتواء جميع المتغيرات التوضيحية، ومن هنا كان لا بُد من تضمين الأنموذج فضلاً عن المتغيرات التوضيحية حداً يمثل البواقي (residuals) وهو الذي يقيس الجزء من التغير في المتغير المعتمد والذي سببتهُ المتغيرات التوضيحية غير المدرجة في أنموذج الانحدار. إن صيغة الانحدار الخطي المتعدد هي:[10] و[7]

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j X_{ij} + \varepsilon_i \quad ...(1)$$

إذ أن: عدد المشاهدات (الأفراد)= n , ... ,2 ,1=i ; عدد المتغيرات التوضيحية= m ,..., 2, 1=j.

y_i: المتغير المعتمد، تتحدد قيمته أساساً عن طريق معالم الأنموذج وقيم المتغيرات التوضيحية والمتغير العشوائي، لذا يُسمى بالمتغير الداخلي.

X_{ij}: قيمة المشاهدة i في المتغير التوضيحي j.

β_o: استجابة المتغير المعتمد عندما لا يوجد تأثير لأيٍ من المتغيرات التوضيحية عليه، وبيانياً هو محل تقاطع خط أو مستوى الانحدار مع المحور الصادي، فعندما تكون قيمة β_o تساوي الصفر (عدم احتواء الأنموذج عليها) عندئذ يمر خط أو مستوى الانحدار من نقطة الأصل.

β_j: معلمات الانحدار (regression Coefficients) التي تُشير قيمها إلى مقدار التغير الذي يحدث في المتغير المعتمد عندما تتغير قيمة المتغير التوضيحي وحدة واحدة، في حالة وجود قيمة سالبة لأي من معلمات الانحدار يعني ذلك وجود علاقة عكسية (سالبة) بين المتغير التوضيحي مع المتغير المعتمد بينما وجود الإشارة الموجبة يُشير إلى وجود علاقة طردية.

ε_i: البواقي (residuals): يمثل حد الخطأ في المشاهدة i (خطأ استجابة المتغير المعتمد).

يمكن باستخدام المصفوفات لصياغة الأنموذج العام بالصورة الآتية:

$$\underline{\hat{\beta}} = (X'X)^{-1}(X'\underline{y}) \quad ...(2)$$

$\underline{y}$: مُتجه عمودي أبعادهُ (n×1) يحتوي مشاهدات المتغير المعتمد.

$$\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_n \end{bmatrix}_{n\times 1} \quad ...(3)$$

$\underline{\mathbf{X}}$: هي مصفوفة أبعادها (n×p) وتحتوي على مشاهدات المتغيرات التوضيحية ويحتوي عمودها الأول على قيم الواحد الصحيح ليُمثل المتغير X_0 المضروب بالحد الثابت β_o الذي قيمتهُ تساوي واحداً:

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}_{n\times p} \quad ...(4)$$

m+1=P

$\underline{\hat{\beta}}$: مُتجه عمودي أبعادهُ (P×1) ويحتوي على معالم الأنموذج المجهولة والمطلوب تقديرها.

$$\underline{\hat{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \vdots \\ \hat{\beta}_m \end{bmatrix}_{p\times 1} \quad \ldots(5)$$

اختيار أفضل معادلة انحدار

إن تقليل عدد المتغيرات التوضيحية التي تُستَخدم في الأنموذج النهائي يَعمَل على التقليل من الجهد والوقت والكلفة وكذلك يضمن سهولة التحليل والفهم، وعلية وجب أن يكون هناك موازنة بين عملية تقليل المتغيرات التوضيحية وزيادة عددها للحصول على نتائج تنبؤيه دقيقة، فيكون من الأفضل اختيار المعادلة المُفسِرة بأقل عدد من المتغيرات التوضيحية بحيث تكون هذه المتغيرات مُهِمة ولها تأثير على المتغير المعتمد[2]. هنالك العديد من طرائق اختيار أفضل أنموذج انحدار فقد أشار Castle et al. (2009) [8] إلى 21 طريقة مختلفة، كما طرح محمود (2011) [6] طريقة مقترحة لاختيار أفضل أنموذج انحدار بالاعتماد على معامل الارتباط الجزئي. أدناه أكثر الطرائق استخداماً وقد يكون سبب ذلك احتواء البرامج والحزم الحاسوبية (packages) عليها.

1- طريقة الاختيار الأمامي (المباشر) Forward selection method

2- طريقة الحذف الخلفي (العكسي) Backward elimination method

3- طريقة الانحدار المتدرج Stepwise regression method

4- طريقة كل الانحدارات الممكنة All Possible regressions method

في طريقة الاختيار الأمامي، الحذف الخلفي وطريقة الانحدار التدريجي سيتم المفاضلة عن طريق اختبار F للفرضيات حول مصادر التباين الكلية والجزئية، أما في طريقة كل الانحدارات الممكنة تتم المفاضلة عن طريقة المقاييس: متوسط مربعات الخطأ MSe وإحصائية مالو Cp [4].

متوسط مربعات الخطأ Mean squares error: يُحدد دقة تأثير المتغيرات التوضيحية، ومع ادلته:

$$MSe = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n-(m+1)} \quad \ldots(6)$$

إنَّ المعادلة ذات أصغر MSe هي التي تُفَضل بأقل عدد من المتغيرات التوضيحية الداخلة في الأنموذج وعليهِ يمكن استخدامهُ معياراً للحكم على ملاءمة المعادلة التقديرية [6].

إحصائية مالو Cp-Mallow Statistics

تعتمد قيمة هذا المقياس على متوسط مجموع مربعات الأخطاء لقيم البيانات الداخلة في الأنموذج الخطي المتعدد الذي يضم المتغيرات كافة (m) المؤثرة على المتغير المعتمد وكذلك المتغيرات للأنموذج المدروس (k)، حيث أن k<m، عن طريق هذا المقياس يتم تحديد المعادلة الأفضل باختيار أقل قيمة Cp، ويمكن إيجاد قيمته وفق الصيغة الآتية:

$$Cp = \frac{SSe(X_1, X_2, \ldots, X_k)}{MSe(X1, X_2, \ldots, X_m)} - (n - 2p) \quad \ldots(7)$$

طريقة الاختيار الأمامي (المباشر) Forward selection method

تبدأ الطريقة باختيار المتغيرات التوضيحية التي تدخل في المعادلة متغيراً بعد الآخر وذلك عن طريق المختبر الإحصائي–F للفرضية الجزئية.

طريقة الحذف الخلفي (العكسي) Backward elimination method

يتم إدخال جميع المتغيرات التوضيحية في معادلة الانحدار كخطوة أولى، وعن طريقة المختبر الإحصائي–F للفرضية الجزئية يتم حذف المتغيرات التوضيحية التي ليس لها دلالة إحصائية الواحد تلو الآخر لحين الوصول إلى الصيغة النهائية للمعادلة التي تحتوي على المتغيرات ذات التأثير المعنوي.

طريقة الانحدار المتدرج stepwise regression method

تسمى أيضاً بطريقة الخطوات المتسلسلة، تدمج بين طريقة الاختيار الأمامي لإدخال المتغيرات إلى الأنموذج وبين طريقة الحذف العكسي التي تعمل على حذف المتغيرات التي ليس لها دلالة إحصائية، ولهذا السبب لها قيمتان جدوليتان إحداها F_{in} والأخرى F_{OUT} [8].

طريقة كل الانحدارات الممكنة All possible regressions method

تعتمد على إيجاد كل النماذج الممكنة (عددها يساوي 2^m) [4] التي تحتوي على متغير توضيحي واحد وحتى عدد m من المتغيرات التوضيحية، توضع النماذج التي تحتوي على نفس عدد المتغيرات التوضيحية في مجاميع عليه فإن عدد المجاميع هو m، وقد يتم إيجاد بعض تلك النماذج وفق ما يراه الباحث ضرورياً، والتي تعطي للمُحلل أكبر كمية من المعلومات المتعلقة بطبيعة العلاقة بين المتغير المعتمد ومجموعة المتغيرات التوضيحية. لاختيار الأنموذج المنشود تتميز هذه الطريقة بالاعتماد على الخبرة والاستعانة بنتائج تحليلية ذات العلاقة وطبيعة البيانات وبعض مقاييس المفاضلة الإحصائية، مثل مجموع مربعات الخطأ MSe وإحصائية مالو Cp.

جمع البيانات

تم الحصول على بيانات ستين متقدماً لامتحان IC3 [4]. جمعت البيانات من مركز الحاسوب والإنترنت/ وحدة الـ IC3 ووحدة الدورات وشَملت الكليات: التربية الأساسية، طب الأسنان، الفنون الجميلة، الهندسة، علوم الحاسوب والرياضيات وكلية التربية. تتكون متغيرات البحث من متغير معتمد وسبعة متغيرات وهي:

Y_i: درجة امتحان التطبيقات الأساسية، مكونةٌ من (1000) درجة، علامة النجاح (750) درجة، وأن عدد الأسئلة هو (45) سؤالاً وبوقت مُحدد مقدارهُ (45) دقيقة.

X_{i1} : درجة امتحان أساسيات الحاسوب.

X_{i2} : جنس المُمتحن (ذكر، أنثى).

X_{i3} : خبرة المُمتحن بالحاسوب وفق الدراسة الجامعية.

X_{i4} : تَخصُص المُمتحن حسب الدراسة في الجامعة (علمي، إنساني).

X_{i5} : مشاركة المُمتحن بالدورات المقامة للـ IC3.

X_{i6} : الشهادة المطلوب الحصول عليها من قِبل المُمتحن (دبلوم عالٍ، ماجستير، دكتوراه).

X_{i7} : إعادة الامتحان (أعاد مرة واحدة، مرتين، ثلاث مرات).

استُخدمت الحزمة البرمجية (Minitab v.13.20) في تحليل البيانات.

إن بيانات المتغير المعتمد (Y_i) والمتغير التوضيحي الأول (X_{i1}) هي بيانات كمية، أما الستة متغيرات التوضيحية الأخيرة فقد رمزت كما في الجدول (1).

الجدول 1. ترميز تشعبات الستة متغيرات التوضيحية الأخيرة وأعداد أفرادها.

المتغير	التشعبات	الرمز	عدد الأفراد
X_2	الذكور	0	31
	الإناث	1	29
X_3	كُليات تُدرس مناهج الحاسوب فيها قليل	1	38
	كُليات تُدرس مناهج الحاسوب فيها متوسط	2	12
	كُليات مختصة بتدريس مناهج الحاسوب	3	10
X_4	التخصص العلمي	0	33
	التخصص الإنساني	1	27
X_5	غير مشارك في الدورات المقامة	0	30
	مشارك في الدورات المقامة	1	30
X_6	دبلوم عالٍ	1	5
	ماجستير	2	34
	دكتوراه	3	21
X_7	لم يُعِدْ الامتحان	0	29
	أعاد مرة واحدة	1	14
	أعاد مرتين	2	10
	أعاد ثلاث مرات	3	7

طبق تحليل الانحدار للتوصل إلى مكونات المعادلة (1)، كما تم تطبيق اختبار F الجزئي[3] وفق المسار الحاسوبي التالي:

stat>Regression>Regression>Response:c1;Predictors:c2-c8>fit intercept → OK.

وقد تم الحصول على النتائج في الجدول (2).

الجدول 2. نتائج تحليل الانحدار بوجود الثابت β_o .

```
The regression equation is
y = 117 +0.776 x1 -58.8 x2 -3.3 x3 -86.6 x4 -63.1 x5 +48.8x6 -22.5 x7

Predictor        Coef     SE Coef          T       P
Constant        116.7       128.7       0.91   0.369
x1             0.7757      0.1298       5.98   0.000
x2             -58.76       30.23      -1.94   0.057
x3              -3.33       24.68      -0.13   0.893
x4             -86.62       39.11      -2.21   0.031
x5             -63.10       32.41      -1.95   0.057
x6              48.75       24.93       1.96   0.056
x7             -22.50       13.40      -1.68   0.099
S = 108.8       R-Sq = 68.6%     R-Sq(adj) = 64.3%
Analysis of Variance
Source            DF        SS           MS        F        P
Regression         7   1342798       191828    16.20    0.000
Residual Error    52    615632        11839
Total             59   1958431
```

يلاحظ في الجداول (2) أن وجود الثابت β_0 في الأنموذج ليس بذي دلالة إحصائية، فضلاً عن أن طبيعة المتغيرات التوضيحية في علاقتها مع المتغير المعتمد قد تُحتم عدم احتواء الأنموذج على الثابت β_0، وهذا يحتم عدم الأخذ بنظر الاعتبار باقي محتويات الجدول (2) حتى يجرى تحليل الانحدار بعدم احتواء الأنموذج على الثابت β_0 أي مرور مستوى الانحدار من خلال نقطة الأصل.

تم إجراء تحليل الانحدار عند عدم احتواء الأنموذج على الثابت β_0 بذات المسار المشار إليه مع إهمال الأمر fit intercept، وقد تم الحصول على نتائج تحليل الانحدار كما في الجدول (*3*).

الجدول (3). نتائج تحليل الانحدار عند عدم احتواء الأنموذج على .β.

```
The regression equation is
y = 0.861 x1 - 56.6 x2 + 5.0 x3 - 64.7 x4 - 48.4 x5 + 60.1 x6 - 21.9 x7

Predictor        Coef     SE Coef          T       P
No constant
x1            0.86069     0.08966       9.60   0.000
x2             -56.60       30.09      -1.88   0.065
x3               5.00       22.86       0.22   0.828
x4             -64.72       30.71      -2.11   0.040
x5             -48.38       28.00      -1.73   0.090
x6              60.12       21.51       2.79   0.007
x7             -21.90       13.36      -1.64   0.107

S = 108.6
Analysis of Variance

Source            DF        SS           MS         F        P
Regression         7   25256816      3608117    305.79    0.000
Residual Error    53     625361        11799
Total             60   25882177
```

يلاحظ من النتائج في الجدول (3) من خلال قيمة F أن الأنموذج المشار إليه ذو دلالة إحصائية عند مستوى معنوية أقل من 0.05 وهذا يدل على أن كل أو بعض المتغيرات التوضيحية تفسر التغيرات الحاصلة في المتغير المعتمد. عليهِ يجب إتباع الطرائق الأربعة في اختيار أفضل أنموذج انحدار عند عدم احتواء الأنموذج على ثابت الانحدار β_0 والاستعانة بمقاييس المفاضلة: قيمة إحصائية مالو Cp ومتوسط مربعات الخطأ MSe.

طريقة الاختيار الأمامي

المسار التالي يوضح نتيجة هذه الطريقة:

Stat>Regression>Stepwise>Response:c1; predictors:c2-c8;Methods>Foreward selection:Alpha to enter:0.05;Options: number of alternative predictors to show: 0; number of steps between pauses:7 → OK.

يبين الجدول (4) المتغيرات التوضيحية ذات الدلالة الإحصائية في تأثيرها على المتغير المعتمد وهي: X_1, X_4, X_5, X_6 ، كما تلاحظ القيمتين: MSe=112 والمتمثلة بـ S في الجدول Cp= 7.3.

الجدول (4). نتائج طريقة الاختيار الأمامي.

When Alpha-to-Enter: 0.05				
Response is y on 7 predictors, with N = 60				
Step	1	2	3	4
No constant				
x1	0.929	0.979	0.810	0.815
T-Value	39.79	33.42	10.47	10.88
P-Value	0.000	0.000	0.000	0.000
x5		-75	-78	-69
T-Value		-2.60	-2.82	-2.53
P-Value		0.012	0.007	0.014
x6			53	60
T-Value			2.35	2.73
P-Value			0.022	0.009
x4				-60
T-Value				-2.18
P-Value				0.033
S	126	120	115	112
C-p	20.8	14.6	10.3	7.3

طريقة الحذف العكسي

يتم تطبيق هذه الطريقة بإتباع المسار الآتي:

Stat> Regression> Stepwise> Response:c1; predictors:c2-c8; Methods> Backward elimination:Alpha to enter:0.05; Options: number of alternative predictors to show: 0; number of steps between pauses:7 → OK.

والجدول (5) يوضح نتائج هذه الطريقة حيث يتبين أن المتغيرات التوضيحية ذات الدلالة الإحصائية في تأثيرها على المتغير المعتمد هي: X_1, X_2, X_4, X_6 ، وبالمقارنة مع نتائج طريقة الاختيار الأمامي يلاحظ أن كليهما اختارا أربعة متغيرات لتكون في النموذج، ولكنهما اختلفا في أهمية المتغيرين X5 و X2 على التوالي، فضلاً عن عدم تغير قيمة متوسط مربعات الخطأ MSe=112 وقد انخفضت قيمة إحصائية مالو لتصبح Cp=7.1، إن هذا الانخفاض يعطي أفضلية لطريقة الحذف العكسي، مع ذلك يجب أن نحصل على نتائج الطرائق الأخرى ومن ثم المفاضلة في اختيار المتغيرات التوضيحية.

الجدول (5). نتائج طريقة الحذف العكسي.

When Alpha-to-Remove: 0.05				
Response is y on 7 predictors, with N = 60				
Step	1	2	3	4
No constant				

x1	0.861	0.871	0.845	0.832
T-Value	9.60	11.53	11.27	10.91
P-Value	0.000	0.000	0.000	0.000
x2	-57	-55	-55	-73
T-Value	-1.88	-1.90	-1.88	-2.56
P-Value	0.065	0.063	0.065	0.013
x3	5			
T-Value	0.22			
P-Value	0.828			
x4	-65	-68	-68	-78
T-Value	-2.11	-2.51	-2.49	-2.85
P-Value	0.040	0.015	0.016	0.006
x5	-48	-48	-52	
T-Value	-1.73	-1.74	-1.84	
P-Value	0.090	0.088	0.072	
x6	60	61	60	59
T-Value	2.79	2.85	2.80	2.70
P-Value	0.007	0.006	0.007	0.009
x7	-22	-22		
T-Value	-1.64	-1.65		
P-Value	0.107	0.106		
S	109	108	109	112
C-p	7.0	5.0	5.7	7.1

طريقة الانحدار المتدرج

بإتباع المسار المشار إليه في طريقة الاختيار الأمامي مع الأخذ بنظر الاعتبار التحوير الآتي:

…Methods> Stepwise (forward and backward); alpha to enter:0.05; Alpha to remove:0.05**…**

يوضح الجدول (6) نتائج هذه الطريقة حيث يبين أن المتغيرات التوضيحية ذات الدلالة الإحصائية في تأثيرها على المتغير المعتمد هي: X_1, X_2, X_4, X_6، وهي ذات المتغيرات التي حُددت بطريقة الحذف العكسي، فضلاً عن أن لهما قيم مقياسي المفاضلة نفسها: MSe=112 و Cp=7.1.

الجدول (6). نتائج طريقة الانحدار المتدرج.

When Alpha-to-Enter: 0.05 and Alpha-to-Remove: 0.05
Response is y on 7 predictors, with N = 60

Step	1	2	3	4
No constant				
x1	0.861	0.871	0.845	0.832
T-Value	9.60	11.53	11.27	10.91
P-Value	0.000	0.000	0.000	0.000
x2	-57	-55	-55	-73
T-Value	-1.88	-1.90	-1.88	-2.56
P-Value	0.065	0.063	0.065	0.013
x3	5			
T-Value	0.22			
P-Value	0.828			
x4	-65	-68	-68	-78
T-Value	-2.11	-2.51	-2.49	-2.85
P-Value	0.040	0.015	0.016	0.006
x5	-48	-48	-52	
T-Value	-1.73	-1.74	-1.84	
P-Value	0.090	0.088	0.072	
x6	60	61	60	59

T-Value	2.79	2.85	2.80	2.70
P-Value	0.007	0.006	0.007	0.009
x7	-22	-22		
T-Value	-1.64	-1.65		
P-Value	0.107	0.106		
S	109	108	109	112
C-p	7.0	5.0	5.7	7.1

طريقة كل الانحدارات الممكنة

يتم التوصل إلى نتيجة هذه الطريقة بإتباع المسار الآتي:

Stat> Regression> Best Subsets> Response:c1; Free predictors:c2-c8; Free Predictor(s) in Each Model: Minimum:1; Maximum:7; Models of each size to print:1 → OK.

طبق المسار للحصول على أفضل سبعة نماذج اختلفت فيما تحتويه من عدد المتغيرات التوضيحية، وتم هذا الاختيار وفق ما أعطاه المسار في الأمر Models of each size to print:1، حيث حصل كل أنمـــوذج على اقل قيمتين لـِ MSe و Cp من بين نماذج كل مجموعة. والنتائج موضحة في الجدول (7).

الجدول (7). نتائج طريقة كل الانحدارات الممكنة.

Variables	C-p	S	x 1	x 2	x 3	x 4	x 5	x 6	x 7
1	20.8	125.54	X						
2	14.6	119.82	X				X		
3	10.3	115.40	X				X	X	
4	7.1	111.62	X	X		X		X	
5	5.7	109.32	X	X		X	X	X	
6	5.0	107.66	X	X		X	X	X	X
7	7.0	108.62	X	X	X	X	X	X	X

تمت مقارنة النتائج في الجدول (7) بين كل النماذج الانحدارية لاختيار الأنموذج الأفضل بأقل قيمتي لإحصائية مالو Cp ومتوسط مجموع المربعات MSe وبأقلِ عدد من المتغيرات التوضيحية التي يعد وجودهـــا مهماً في الأنموذج كونها قد تؤثر بشكل كبير على المتغير المعتمد.

تم تعيين الأنموذج الذي يحتوي على المتغيرات التوضيحية الأربعة: X_6, X_4, X_2, X_1، بأنه أفـــضل أنموذج أعطته هذه الطريقة، حيث يلاحظ توازن بين قيم MSE و Cp وعدد متغيرات الأنموذج، فضلاً عن أن هذه النتائج أعطتها كل من طريقة الحذف العكسي وطريقة الانحدار المتدرج.

مقارنة نتائج الطرائق الأربعة

بعد أن تم اختيار أفضل أنموذج وفق نتائج الطرائق الأربعة، لُخصتْ النتائج في الجدول (8).

الجدول (8). ملخص نتائج الطرائق الأربع لاختيار أفضل أنموذج انحدار.

الطريقة المستخدمة	Cp	MSe	المتغيرات المختارة
الاختيار الأمامي	7.3	112.0	(X_6, X_5, X_4, X_1)
الحذف العكسي	7.1	112.0	(X_6, X_4, X_2, X_1)

الانحدار المتدرج	7.1	112.0	(X_6,X_4,X_2,X_1)
كل الانحدارات الممكنة	7.1	111.62	(X_6,X_4,X_2,X_1)

من الجدول (8) يلاحظ تقارب ما أعطته الطرائق من قيم MSe، واختلاف طفيف لقيمة Cp في طريقة الاختيار الأمامي عن بقية الطرائق والتي اختلفت أيضاً بمتغيرات الأنموذج عن الطرائق الثلاثة الأخرى التي أعطت الأنموذج ذاته، وسيتم اختيار أحد الأنموذجين من خلال دراسة تحليل الانحدار لهما، وقد كانت النتائج في الجدولين (9) و(10) على التوالي.

الجدول (9). نتائج تحليل الانحدار لأنموذج المتغيرات X_6,X_5,X_4,X_1 .

```
The regression equation is
y = 0.815 x1 - 60.4 x4 - 68.7 x5 + 60.1 x6

Predictor        Coef     SE Coef        T        P
Noconstant
x1            0.81532     0.07495    10.88    0.000
x4             -60.40       27.68    -2.18    0.033
x5             -68.69       27.19    -2.53    0.014
x6              60.09       22.04     2.73    0.009

S = 111.8

Analysis of Variance

Source            DF        SS         MS        F        P
Regression         4  25182594    6295648   503.95    0.000
Residual Error    56    699583      12493
Total             60  25882177
```

الجدول (10). نتائج تحليل الانحدار لأنموذج المتغيرات X_6,X_4,X_2,X_1 .

```
The regression equation is
y = 0.832 x1 - 72.9 x2 - 78.2 x4 + 59.5 x6
Predictor        Coef     SE Coef        T        P
Noconstant
x1            0.83205     0.07626    10.91    0.000
x2             -72.86       28.46    -2.56    0.013
x4             -78.18       27.41    -2.85    0.006
x6              59.49       22.01     2.70    0.009
S = 111.6
Analysis of Variance
Source            DF        SS         MS        F        P
Regression         4  25184489    6296122   505.36    0.000
Residual Error    56    697688      12459
Total             60  25882177
```

بالمقارنة بين محتوى الجدولين (9 و10) يلاحظ تقارب بين كافة معطياتهما عدا قيم P للاختبار الجزئي للمتغيرين X_2 وX_4 فهي أصغر في الجدول (10) وهذا يعني قلة احتمالية الوقوع في الخطأ من النوع الأول بإدخالهما إلى الأنموذج معاً، فضلاً عن أن طريقة الحذف العكسي تطابقت مع ما أعطتهُ طريقة الانحدار المتدرج وكل الانحدارات الممكنة في اختيار الأنموذج الذي يتضمن المتغيرات التوضيحية X_6,X_4,X_2,X_1 ، وبهذا يكون أفضل أنموذج انحدار يضم المتغيرات المؤثرة على درجة امتحان التطبيقات الأساسية (المتغير المعتمد y_i) هو:

$$\hat{y}_i = 0.832\ x_{i1} - 72.9\ x_{i2} - 78.2\ x_{i4} + 59.5\ x_{i6} \quad \ldots(8)$$

يتضح من الأنموذج أن درجة امتحان التطبيقات الأساسية (المتغير المعتمد y_i) المكونة من 1000 درجة تتأثر كالآتي: إن زيادة درجة في نتيجة امتحان أساسيات الحاسوب تؤدي إلى زيادة 0.832 درجة في نتيجة امتحان التطبيقات الأساسية، بينما إذا كان المشارك أنثى يؤدي ذلك إلى خفض الدرجة بمقدار 72.9 درجة، وأيضاً التخصصات الإنسانية تعمل على خَفض الدرجة بمقدار 78.2 درجة، في حين أن تأثير الشهادة المطلوب الحصول عليها كان المتقدم لدراسة الماجستير تزيد درجته عن المتقدم لدراسة الدبلوم العالي بمقدار 59.5 درجة، ونفس الزيادة للمتقدم لدراسة الدكتوراه عن المتقدم لدراسة الماجستير.

الاستنتاجات

1- تبين من التطبيق عدم جدوى الأنموذج المتضمن ثابت الانحدار β_0، عليه تم إلغاء الثابت β_0 من الأنموذج ثم طبقت أربع طرائق لاختيار أفضل أنموذج انحدار يفسر التغيرات في درجة امتحان التطبيقات الأساسية.
2- إن طريقة الاختيار الأمامي قد حددت المتغيرات التوضيحية X_6, X_5, X_4, X_1 لتكون في الأنموذج. في حين أن طريقة الحذف العكسي وطريقة الانحدار المتدرج وطريقة كل الانحدارات الممكنة حددت المتغيرات التوضيحية X_6, X_4, X_2, X_1 لتكون في الأنموذج، فضلاً عن إعطائها نتائج تحليلية متطابقة.
3- بعد المقارنة بين نتائج تحليل الانحدار تم اختيار الأنموذج المتضمن المتغيرات التوضيحية : X1: درجة الامتحان في أساسيات الحاسوب وX2: جنس المُمتحن وX4: تخصص الممتحن حسب الدراسة في الجامعة وX6: الشهادة المطلوب الدراسة إليها، ومعادلة الانحدار (8) تمثل الأنموذج المختار.

التوصيات

1- الدراسات التي تحتوي على عدد كبير من المتغيرات التوضيحية والتي يُعتقد أنّ عدداً منها له تأثير على المتغير المعتمد، نوصي بتطبيق كافة الطرائق ومن ثم النظر إلى طبيعة البيانات وعلاقتِها العلمية وكذلك اختبار الفرضيات الجزئية للمتغيرات التوضيحية وإضافة الخبرة العلمية والعملية مع الأخذ بنظر الاعتبار مقاييس المفاضلة، وذلك لبِناء فكرة صائبة لاختيار أفضل أنموذج.
2- دراسة أسباب التأثير السلبي لجنس المُمتحن وتخصص الممتحن حسب الدراسة في الجامعة على درجة امتحان التطبيقات الأساسية.
3- الاهتمام بالمتغيرات التوضيحية التي ليس لها أهمية في تأثيرها على درجة امتحان التطبيقات الأساسية والتعرُّف على أسباب عدم جدواها وهي: خبرة المُمتحن بالحاسوب وفق الدراسة الجامعية ومشاركة المُمتحن بالدورات المقامة للـ IC3 وإعادة الامتحان.

المصادر

[1] الجاعوني، فريد خليل، 2008، "أسلوب تحليل الانحدار الخطي المتعدد في دراسة أهم المتغيرات الاقتصادية والاجتماعية والديمغرافية المؤثرة في معدل الولادات الكلية"، مجلة العلوم الاقتصادية والقانونية، المجلد 24، العدد 2،ص237-253، جامعة دمشق- سوريا.

[2] النعيمي، أسوان محمد طيب ، 2005، "اختيار المتغيرات في انحدار الحرف"، رسالة ماجستير غير منشورة، كلية علوم الحاسوب والرياضيات، جامعة الموصل- العراق.

[3] دبدوب، مروان عبد العزيز وهدبة، وكاع علي، 2008، "العلاقة بين تحليل التباين وتحليل انحدار المتغيرات الوهمية"، المجلة العراقية للعلوم الإحصائية، المجلد 8، العدد 14، ص113-132، كلية علوم الحاسوب والرياضيات، جامعة الموصل- العراق.

[4] زكي، أحمد محمد، 2012، "استخدام عدد من المقاييس الإحصائية الشائعة في اختيار أفضل أنموذج انحدار متعدد بالتطبيق على درجة اختبار IC3"، رسالة دبلوم عالي غير منشورة، كلية علوم الحاسوب والرياضيات، جامعة الموصل- العراق.

[5] شركة الخليج للتدريب والتعليم، 2005، "الشهادة الدولية للحاسب والانترنت"، ط ٣، الرياض- المملكة العربية السعودية.

[6] محمود، غزوان هاني، 2011، "طريقة مقترحة لاختيار أفضل أنموذج معادلة انحدار بالاعتماد على معامل الارتباط الجزئي"، المؤتمر العلمي الرابع 21-2011/12/22، كلية علوم الحاسوب والرياضيات، جامعة الموصل- العراق.

[7] Bluman, A. G.; 2012; "Elementary Statistics: A Step By Step Approach"; 8th Edition; McGraw-Hill Companies, Inc.; USA.

[8] Castle, J.; Qin, X. and Reed, W.; 2009; "How To Pick The Best Regression Equation: A Review And Comparison Of Model Selection Algorithms"; Working Paper; Department of Economics and Finance; College Of Business And Economics; Univ. of Canterbury; New Zealand.

[9] Chatterjee, S. and Hadi, A.; 2006; "Regression Analysis By Example"; 4th Edition; John Wiely and Son's, Inc.; USA.

[10] Yan, X. and Su, X.; 2009; "Linear Regression Analysis Theory And Computing"; Word Scientific Publishing; USA.