جامعة محمد خيضر (بسكرة)
كلية العلوم الاقتصادية والتجارية والتسيير
قسم العلوم التجارية

السنة الجامعية: 2026/2025
المستوى: أولى ماستر تسويق
مقياس: الأساليب الكمية في التسويق 2

أعمال تطبيقية رقم (01) حول الانحدار اللوجستي Logistic regression with SPSS

مصطلحات أساسية

قبول الطالب في الدراسات العليا Admet:

متغير تابع dependent، اسمي nominal، ثنائي Dichotomique يأخذ قيمتين: 0 لا يقبل، 1 يقبل.

اختبار تقييم الخريجين graduate record examinations (gre)

اختبار القدرات أو التحصيل للطلبة، وهو اختبار قياسي متعدد، محوسب، معتمد وشرط للقبول في كليات الدراسات العليا في الولايات المتحدة والبلدان الناطقة بالإنجليزية، يهدف إلى قياس: المنطق الشفهي، والمنطق الكمي، والكتابة التحليلية، ومهارات التفكير الناقد المكتسبة على مدى فترة طويلة من الزمن وغير المرتبطة بأي حقل معين من الدراسة.

المعدل التراكمي grade point average (gpa):

يرمز إلى متوسط الدرجات، ويحسب بقسمة مجموع الدرجات التي حصل عليها الطالب في حياته الدراسية أو المهنية مقسومًا على إجمالي عدد الاعتمادات، وتعلن معظم المدارس الثانوية والكليات عن درجات الطلاب على مقياس يتراوح بين معدل أدنى 0 ومعدل أعلى 4.

المتغيران: القدرات gre والمعدل التراكمي gra هما متغيران كميان قياسيان Echelle، الأول يتراوح بين 0 للأدنى، و800 للأفضل، والثاني يتراوح بين 0 للأدنى و4.00 للأفضل.

Rank رتبة الطالب، متغير فئوي Categoricale variable أو Ordinale variable، يتراوح بين 1 و4، حيث 1 هي أفضل رتبة، و4 هي أسوأ رتبة، وهو ما يعني أن الطالب الذي يحصل على rank يساوي 1 أو 2 يرتفع احتمال قبوله في الدراسات العليا، والعكس صحيح بالنسبة لطالب رتبته 3 أو 4.

لدينا ثلاث متغيرات (gre؛ gra؛ rank) يفترض أنها تؤثر بالسلب أو الإيجاب في قبول الطالب في الدراسات العليا.

فمثلا الطالب رقم 6: لديه gre= 760، أي مستوى القدرات والتحصيل مرتفع، وgpa= 3، أي معدل الدرجات طيلة مشواره الدراسي مرتفع نوعها ما (3 أقرب لـ 4)، كما أن rank= 2، أي لديه رتبة متقدمة، ومنه تم قبوله في الدراسات العليا admet= 1.

مثلا الطالب رقم 5: لديه gre= 520، أي مستوى فوق المتوسط في القدرات والتحصيل، وgpa= 2.93، أي معدل درجاته فوق المتوسط، كما أن rank= 4، أي لديه رتبة متأخرة، ومنه تم رفضه في الدراسات العليا admet= 0.

الخطوة (1): إدخال البيانات الواردة في المثال (انطلاقا من ملف Exel الأول المرفق) إلى نافذة البيانات SPSS (Data View)، ثم في نافذة المتغيرات Variable View، يتم تسمية 4 متغيرات: القبول في الدراسات العليا admet (وهو متغير نوعه: String قياسي، اسمي Nominal ثنائي)، اختبار القدرات gre والمعدل التراكمي gpa (متغيران عدديان Numeric)، والرتبة rank وهو متغير فئوي Cardinale أو رتبي Ordinale. وهنا تظهر صفحة البيانات كما يلي:

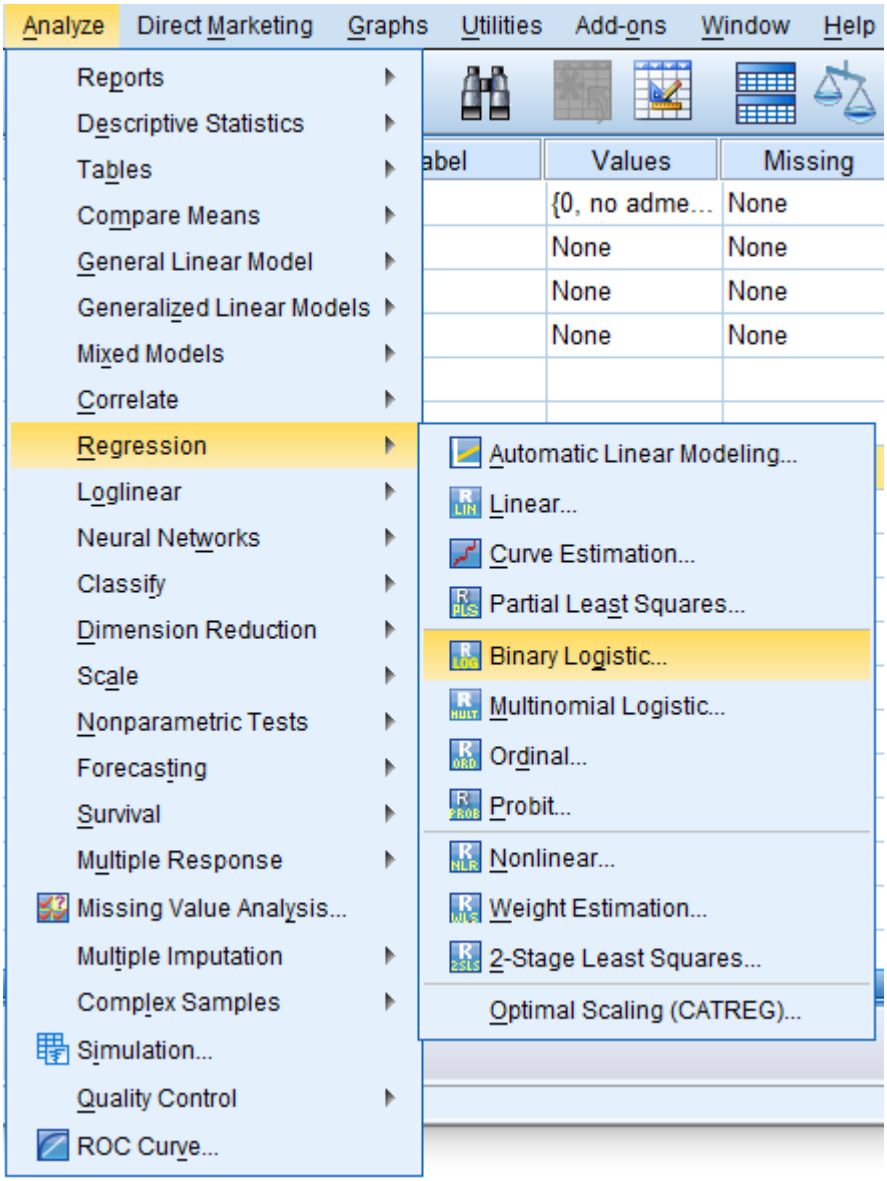
Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
admet	String	1	0		{0, no adme...	None	8	Left	Nominal	Input
gre	Numeric	8	0		None	None	8	Right	Scale	Input
gpa	Numeric	8	2		None	None	8	Right	Scale	Input
rank	Numeric	8	0		None	None	8	Right	Nominal	Input

وبعد تسمية المتغيرات وتحديد خصائصها تظهر نافذة البيانات Data View كما يلي:

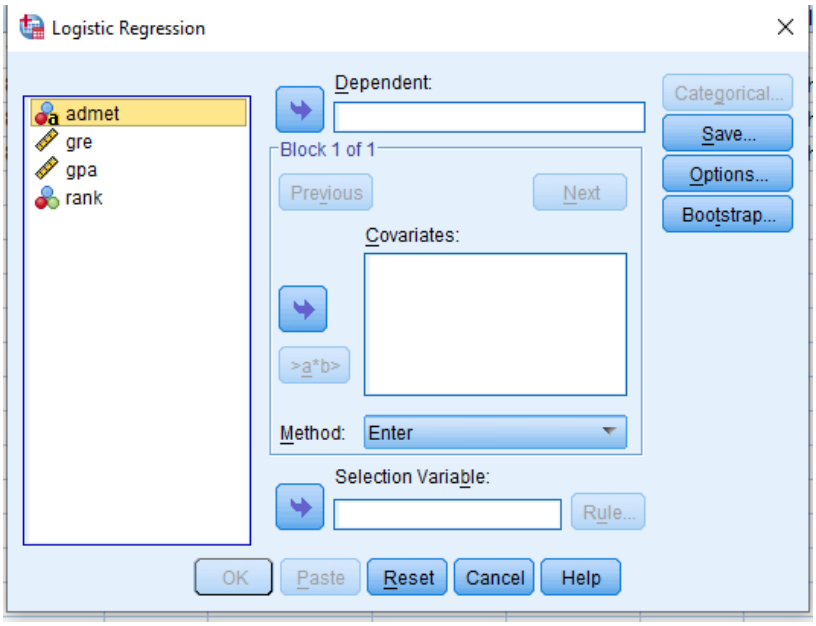
	admet	gre	gpa	rank
1	0	380	3.61	3
2	1	660	3.67	3
3	1	800	4.00	1
4	1	640	3.19	4
5	0	520	2.93	4
6	1	760	3.00	2
7	1	560	2.98	1
8	0	400	3.08	2
9	1	540	3.39	3
10	0	700	3.92	2
11	0	800	4.00	4

الخطوة (2): إجراء التحليل اللوجستي

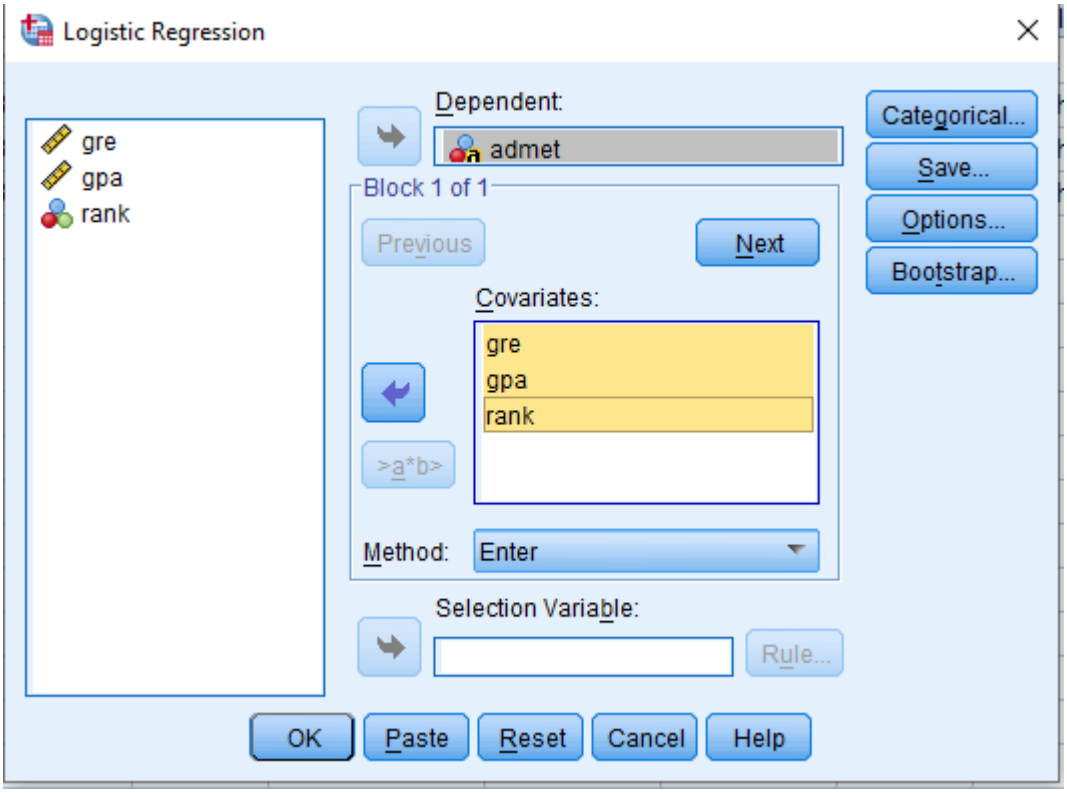
اختيار القائمة الرئيسة Analyze، ثم القائمة لفرعية Regression، ثم الضغط على Binary Logistic، أي اختيار التحليل اللوجستي الثنائي، لأن المتغير التابع admet اسمي ثنائي binomial ، حيث يأخذ قيمتان فقط (0 و 1).



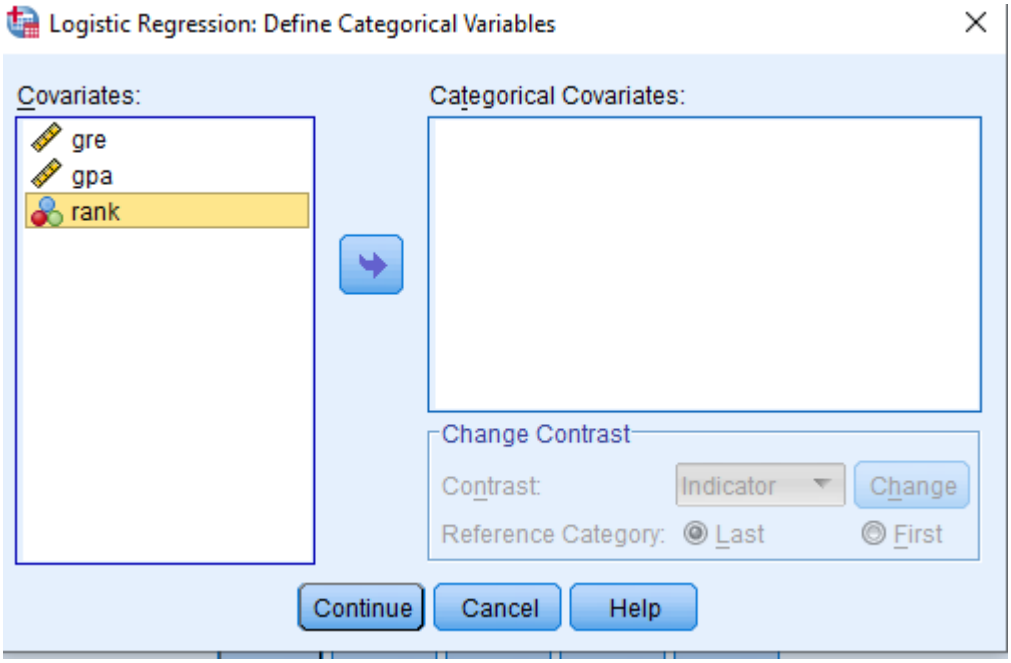
يظهر صندوق الحوار Logistic Regression التالي:



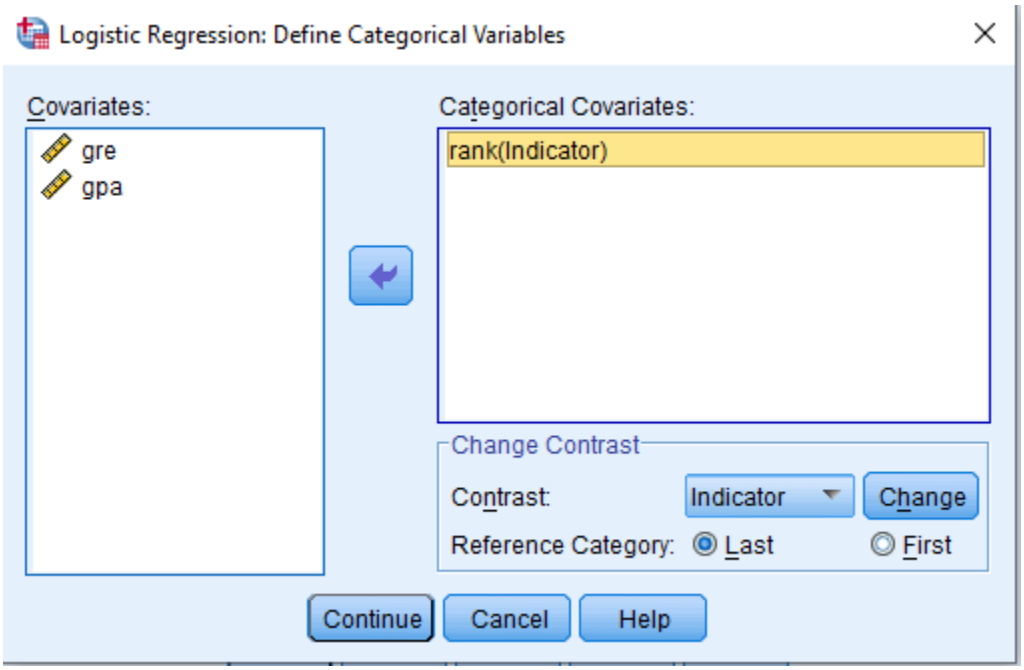
يتم نقل المتغير التابع admet إلى داخل المستطيل المعنون Dependent، ونقل المتغيرات gre وgpa و rank إلى داخل المستطيل المعنون Covariates كما في الشكل التالي:



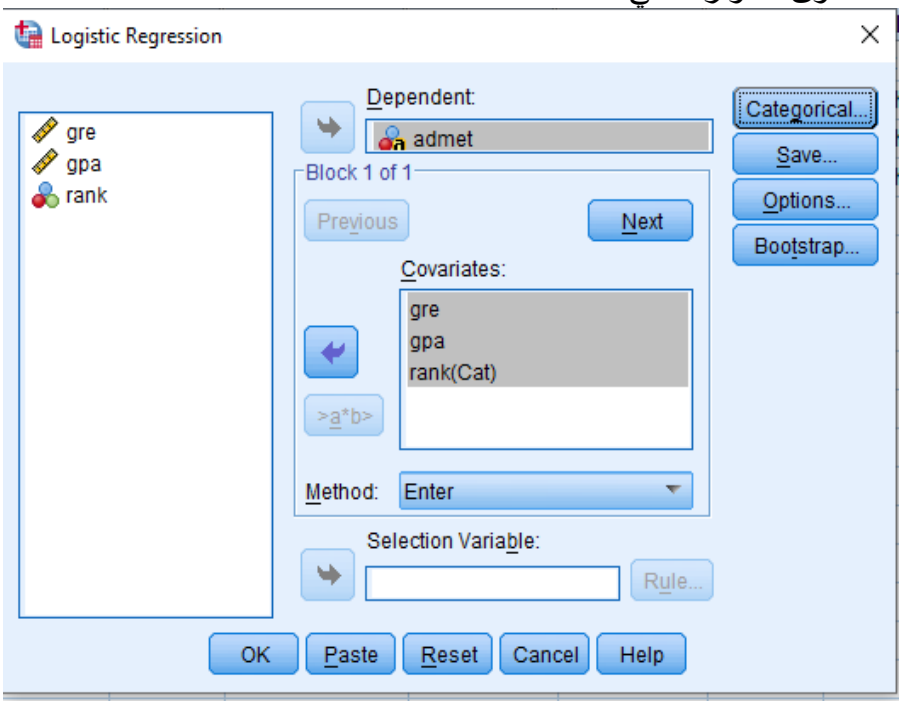
لكن بما أن المتغير rank فئوي، فيجب تعريفه، وهذا حتى نتعرف على فئات هذا المتغير والتي تؤثر أو لا تؤثر على النموذج (قبول أو عدم قبول الطلبة في الدراسات العليا). ولهذا نضغط على Categorical، فيظهر صندوق الحوار التالي:



نقوم بنقل المتغير rank إلى المستطيل Categorical Covarintes، ونضغط على Continue في الأسفل، فيظهر صندوق الحوار التالي:



نضغط على Continue، فيظهر صندوق الحور Logistic Regression، وقد صار المتغير rank متغير فئوي (cat)، على أن نبقي على الطريقة القياسية Enter في خانة الطريقة Method، وenter تعني إدخال المتغيرات المستقلة إلى نموذج الانحدار، حيث يتم إدخال جميع المتغيرات المستقلة في دفعة واحدة دون فحص أي المتغيرات لها أثر ذو دلالة على المتغير التابع admet، وهو ما يوضحه صندوق الحوار التالي:



الخطوة (3): نضغط ok في صندوق الحوار الأخير Logistic Regression، فتظهر **مخرجات التحليل اللوجستي** التالية:

Logistic Regression

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	400	100.0
	Missing Cases	0	.0
	Total	400	100.0
Unselected Cases		0	.0
Total		400	100.0

a. If weight is in effect, see classification table for the total number of cases.

يمثل الجدول السابق عدد الحالات التي دخلت في الاختبار (400 طالب)، وعدد الحالات التي تقابله بيانات مفقودة لا يوجد Missing Cases= 0.

Dependent Variable Encoding

Original Value	Internal Value
no admet	0
admet	1

يمثل الجدول السابق تشفير Codings للمتغير التابع admet، حيث يأخذ قيمتين: 0 تمثل عدم القبول، و1 تمثل القبول في الدراسات العليا.

Categorical Variables Codings

		Frequency	Parameter coding		
			(1)	(2)	(3)
rank	1	61	1.000	.000	.000
	2	151	.000	1.000	.000
	3	121	.000	.000	1.000
	4	67	.000	.000	.000

يمثل الجدول السابق Codings للمتغير الفئوي Categorical Varibles، وهو الرتبة rank، حيث يأخذ أربع قيم صحيحة (من 1 إلى 4)، وتظهر في الجدول تكرارات Frequency هذا المتغير، فمثلا عدد الطلبة الذين حصلوا على الرتبة الأولى (1): 61، وعدد الطلبة الذين حصلول على الرتبة الثانية (2): 151، ... وهكذا،

Block 0: Beginning Block

يمثل Block 0 النموذج الأولي أو الابتدائي للانحدار اللوجستي، النموذج من دوي المتغيرات المفسرة التنبؤية Predictors Variables (gre، gpa، rank)، ولذا لا يمكن لا يمكن اعتباره نموذج إحصائي Nule model وهذا حتى نختبر عند إدخال هذه المتغيرات التنبؤية المفسرة، هل تحقق تحسين في النموذج أم لا. أم أن النموذج اللوجستي مع هذه المتغيرات لا يحث له تحسين. وبتعبير آخر هل يساهم إدخال المتغيرات (gre، gpa، rank) في التنبؤ بشكل صحيح بقبول الطالب بالدراسات العليا من عدمه. أو هل لهذه المتغيرات علاقة بقبول الطالب بالدراسات العليا من عدمه.

Classification Table[a,b]

Observed			Predicted		
			admet		Percentage Correct
			no admet	admet	
Step 0	admet	no admet	274	0	100.0
		admet	126	0	.0
	Overall Percentage				68.5

a. Constant is included in the model.

b. The cut value is .500

في الجدول السابق لا يوجد أي متغير تنبؤي (تفسيري) Predicted من المتغيرات الثلاث: gre، gpa، rank، ولا يوجد إلا متغير واحد ملاحظ Observed، وهو المتغير الملاحظ Observed، حيث لدينا حالتان: غير مقبول no admet (274 طالب)، و مقبول admet (126)، وهذه القيم لم يتم الاعتماد في الحصول عليها على المتغيرات المستقلة المفسرة الثلاث: gre، gpa، rank، ولكن بالاعتماد على قيم المتغير التابع admet في جدول المعطيات، وهو ما أعطاه أيضا جدول التكرارات، ومنه فالمرحلة الأولى تبين أن عدد الطلبة الذين تم قبولهم في الدراسات العليا 126، بنسبة إحصائية شاملة Overall Percentage تساوي 68.5%.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-.777	.108	52.087	1	.000	.460

يمثل الجدول السابق أن المتغيرات المفسرة gre، gpa، rank، إذا تم إدخالها في معادلة الانحدار اللوجستي، فإنها ستساهم مجتمعة في التنبؤ بقبول أو عدم قبول الطلبة في الدراسات العليا، حيث أن اختبار Wald يساوي 52.087، وهو

ذو دلالة المعنوية (Sig أقل من 0.05)، أما دالة اللوجيت *Logit* فهي ذات معامل انحدار بيتا= -0.777، أي أن له قيمة أسية 0.460 =$e^{-0.777}$= Exp (-.777)، والخطأ المعياري في قياسه يساوي: SE= 0.108.

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	gre	11.707	1	.001
		gpa	11.720	1	.001
		rank	24.997	3	.000
		rank(1)	17.035	1	.000
		rank(2)	1.456	1	.228
		rank(3)	5.618	1	.018
	Overall Statistics		38.393	5	.000

الجدول السابق هو أهم جدول في نموذج الانحدار اللوجستي، حيث يوضح المتغيرات التي استبعدت من النموذج وأثرها، أي هل لو تم إضافتها ستؤثر على النموذج أم لا، بعبارة معاكسة، هل لو تم إضافتها إلى المعادلة، فإن ذلك قد يسمح بتحسين النموذج (التنبؤ بقبول أو عدم قبول الطالبة في الدراسات العليا).

نلاحظ أن مستوى الدلالة المعنوية Sig للمتغيرات المفسرة الثلاث (gre، gpa، rank) أقل من 0.05 (5%)، وهو ما يؤكد أن إدخاله في النموذج (معالة الانحدار اللوجستي) سيساهم في تحسين النموذج (ارتفاع القدرة على التنبؤ بقبول أو عدم قبول الطلبة في الدراسات العليا).

رغم أن المتغير المفسر rank يساهم إدخاله في معادلة الانحدار اللوجستي في تحديد ما إذا كان الطالب سيقبل أو لا في الدراسات العليا، إلا أن إحدى قيم هذل المتغير لا تساهم في ذلك، وهي القيم الرتبة 2 ، حيث يقابلها في الجدول السابق مستوى معنوية Sig= 0.228، وهي أكبر من 0.05 ، أما الرتبتين 1 و2 فتساهم في تحسين النموذج (التنبؤ بقبول الطالب في الدراسات العليا وبشكل صحيح باحتمال كبير)، حيث مستوى معنوية الرتبة rank عند 1 و3 أقل من 0.05.

Block 1: Method = Enter

يمثل Block 1 نموذج الانحدار اللوجستي بعد إدخال المتغيرات المفسرة (gre، gpa، rank) بطريقة الانحدار القياسي Enter، وهو يسمح بتقدير قيمة بيتا، وتحديد هل ادخال المتغيرات الداخلة له تأثير في التنبؤ بالقبول في الدراسات العليا أم لا. أما Enter فتشير إلى اختيار الطريقة القياسية في تحليل الانحدار.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	459.118[a]	.094	.131

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

يعرض الجدول السابق بعض المؤشرات الإحصائية المعروفة لتقييم جودة نموذج الانحدار اللوجستي، أي مدى قدرة النموذج الرياضي المستخرج على إعطاء قيم تنبؤية Predicted للمتغير التابع admet تكون أقرب للقيم الملاحظة Observed المعطاة في جدول المعطيات في المثال.

تعد نسبة الإمكان الأعظم Likelihood Ratio الذي يرمز له ب LR(k)، والذي تتبع توزيعكاي تربيع Ch-2، أهم اختبارات كفاية النموذج بالكامل وجودته Goodness of fit ، حيث كان يعتمد لتحقيق هذا الهدف في الانحدار الخطي على اختبار فيشر F أو Student -T، بينما يتم استخدام اختبار نسبة لوغاريتم الإمكان الأعظم Log Likelihood Ratio، ولأغراض رياضية يستخدم -2 Log، وذلك لتحديد معنوية النموذج إحصائيا بالنسبة لأي متغير مستقل، وذلك من خلال اختبار $LR(k)$ للنموذج الذي يحتوى ذلك المتغير مقارنة مع النموذج الذي لا يحتوى ذلك المتغير المستقل.

نلاحظ انخفاض نسبة -2 Log likelihood من القيمة 498.55 التي أعطاه الجدول Iteration History في النموذج الأولي في Block 0، في حالة عدم إدخال المتغيرات التفسيرية، إلى القيمة 459.12، في حالة إدخال المتغيرات التفسيرية، وهو ما يشير إلى تقلص في التباين بين القيم الملاحظة في جدول المعطيات، والقيم التنبؤية من النموذج بـ 39.43، وهو تحسن في النموذج بعد إدخال متغيرات gre، gpa و rank.

ومن الجدول السابق، نجد أن المتغيرات التفسيرية تساهم في التنبؤ بقبول أو عدم قبول الطلبة بنسبة 9.40 %، وهي نسبة ضعيفة، لأن قيمة Cox & Snell R^2 تساوي قيمة 0,094، وبما أن الحد الأقصى لهذه النسبة هو 0.75، فقد تم تعديلها بنسبة تسمى Nagelkerke R^2 المعدلة، والتي تنحصر بين 0 و1، وبما أنها تساوي 0.131، فهذا يعني أن نسبة

%13,10 فقط من التغيرات في المتغير التابع الثنائي (admet) تفسرها المتغيرات المستقلة (gre, gpa, rank). وهي نسبة تفسيرية جد ضعيفة.

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	39.316	5	.000
	Block	39.316	5	.000
	Model	39.316	5	.000

تستخدم الاختبارات الشاملة Omnibus Tests لمعاملات النموذج للتحقق من أن النموذج الجديد (بما في ذلك المتغيرات التفسيرية: Block 1)، يمثل تحسنًا عن النموذج الأول(بدون متغيرات تفسيرية: Block 0). وهنا يستخدم اختبار Chi-2 لمعرفة ما إذا كان هناك فرق بين لوغاريتم الإمكان الأعظم Log-likelihoods (Log LR)، وبشكل خاص: -2Log(LR) ، للنموذج الأولي والنموذج الجديد بعد إدخال المتغيرات التفسيرية.

نلاحظ أن اختبار Ch-2 ذو دلالة معنوية (Sig < 0.05)، وهو منخفض جدا مقارنة بالنموذج الأولي (انتقل من 498.55 إلى 39.31 فقط)، مما يدل على تحسن كبير جدا في التنبؤ بقبول الطلبة في الدراسات العليا عن طريق استخدام متغيرات gre ، gpa وrank، وهو ما يعني أن نموذج الانحدار اللوجستي المستخرج أفضل بشكل كبير في التنبؤ بقبلول الطلبة في الدراسات العليا.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	8.351	8	.400

يستخدم اختبار Hosmer & Lemeshow (H&L) للتأكد فيما إذا كان نموذج الانحدار اللوجستي يمثل المعطيات الفعلية بشكل جيد أم لا، وذلك عن طريق استخدام اختبار Ch-2 لحسن المطابقة لتقييم الفرق بين القيم المشاهدة (Observed) التي لا تستند إلى نموذج نظري، والقيم المتوقعة (Expected) التي تم احتسابها من تقديرات النموذج المقدر.

تشير إحصائية (H&L) إلى ملاءمة النموذج إلى البيانات المشاهدة، إذا كانت مستوى معنويته أقل من0.05. هنا نموذج الانحدار اللوجستي التنبؤي يلائم البيانات المشاهدة الفعلية بشكل كا ف، ويعتبر (H&L) الاختبار الأكثر موثوقية لملاءمة النموذج للانحدار اللوجستي الثنائي، لأنه يجمع المشاهدات في مجموعات من الحالات" المتشابهة."

نلاحظ من الجدول السابق أن اختبار (H&L) ليس ذو دلالة معنوية (Sig= 0.400< 0.05)، بما يدل على ضعف ملاءمة أو جودة نموذج الانحدار اللوجستي، أي ضعف قدرته في التنبؤ بقبول /عدم قبول الطلبة في الدراسات العليا.

Classification Table[a]

Observed			Predicted		
			admet		Percentage Correct
			no admet	admet	
Step 1	admet	no admet	256	18	93.4
		admet	96	30	23.8
	Overall Percentage				71.5

a. The cut value is .500

نلاحظ في جدول التصنيف أعلاه الذي يقارن بين المشاهدات الفعلية (Observed) والتوقعات (Predicted) أعطاها نموذج الانحدار اللوجستي، أنه تم التوافق بينهما في: 256 + 30 = 286 طالب، بنسبة إجمالية Overall Percentage وفق نموذج الانحدار اللوجستي: 71.5%، وعدم التوافق في: 114 = 18 + 96 طالب. أما النسب الصحيحة انطلاقا من جدول التمرين فهي: 256÷ (256+18) = 93.4% طالب غير مقبول، و: 30 ÷(30 + 96)= 23.8 % طالب مقبول في الدراسات العليا.

بالمقارنة بين النموذج قبل وبعد إدراج المتغيرات التفسيرية، نجد أن نسبة التنبؤ السليمة بالالتحاق /عدم الالتحاق زادت من %68.5 (274 طالب)، إلى %71.5 (256+30 = 286 طالب)، أي أن التحسن في التنبؤ باستخدام نموذج الانحدار اللوجستي كان ضعيف 3% = 68.5 - 71.5 (فقط)

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
	gre	.002	.001	3.512	1	.061	1.002
	gpa	.790	.334	5.610	1	.018	2.204
	rank			20.950	3	.000	
Step 1[a]	rank(1)	1.573	.417	14.247	1	.000	4.821
	rank(2)	.861	.366	5.531	1	.019	2.366
	rank(3)	.232	.392	.351	1	.554	1.261
	Constant	-5.384	1.144	22.161	1	.000	.005

a. Variable(s) entered on step 1: gre, gpa, rank.

يتضح من الجدول السابق أن المتغيرات التي يقابلها مستوى معنوية أقل من 0.05 (5%)، تساهم في تحسين نموذج الانحدار اللوجستي، أي تساهم في تفسير والتنبؤ بقبول الطالب في الدراسات العليا، وهي المتغيرات: gpa، وrank، أما متغير gre فلا يساهم كثيرا في التنبؤ بالقبول في الدراسات العليا، لأن مستوى معنويته تساوي 0.61، وهو أكبر من 0.05.

أما مستويات الرتبة، فإن الـ rank 1 و2، فإنه لهما تأثير على قبول الطالب في الدراسات العليا، وهذا لأن مستوى المعنوية Sig عنده أقل من 0.05، في حين أن الرتبة 3 ليس لها تأثير على قبول الطالب في الدراسات العليا، لأن مستوى المعنوية عندها 0.554 أكبر من 0.05.

يمكن كتابة معادلة خط الانحدار اللوجستي كما يلي: $odds= e^{-5,384 + 0,790\ gpa\ + 1,573\ rank\ (1) + 0,861\ rank\ (2)}$

وبإدخال اللوغاريم على دالة الأرجحية، نحصل على نموذج الانحدار اللوجستي مع المتغيرات الدالة احصائيا فقط:

$$Log\ (odds)= -5,384 + 0,790\ gpa + 1,573\ rank\ (1) + 0,861\ rank\ (2)$$