

الانحدار اللوجستي

Logistic Regression

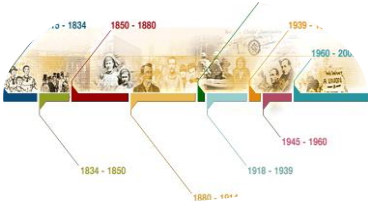
د. سيف بن فهد القحطاني

إحصاء نفسي 605

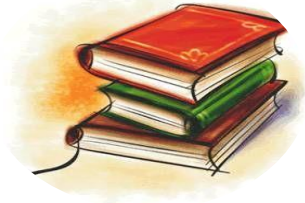
أكتوبر 2015



الانحدار اللوجستي
Logistic Regression



المفاهيم الأساسية



الشروط



مثال



SPSS

الانحدار اللوجستي

Logistic Regression

- الفكرة الرئيسة
- الشروط
- إدخال البيانات
- التحليل
- النتائج
- التفسير

الانحدار اللوجستي

Logistic Regression

□ الفكرة الرئيسة

- الكشف عن أثر متغير مستقل (أو أكثر) في متغير تابع
- المتغير التابع على المستوى الاسمي التصنيفي ثنائي المستويات (نعم-لا)

● الفكرة الرئيسة

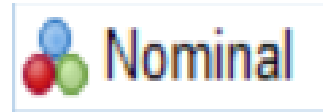
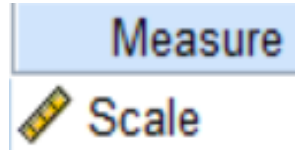
- التنبؤ بمتغير تصنيفي (ثنائي المستويات - ناجح / راسب) من خلال متغير مستقل أو أكثر

الانحدار اللوجستي

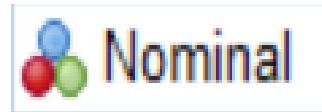
Logistic Regression

• الشروط

• المتغيرات المستقلة يمكن أن تكون على المستوى الاسمي التصنيفي أو الكمي



• المتغير التابع على المستوى النوعي (الاسمي التصنيفي) الثنائي



• المتغير التابع (عدد 1) / المتغير المستقل (عدد 1 أو أكثر)

الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

الأرجحية (Odds)

$$\frac{p}{q}$$



دعنا نفترض أنك ترمي قطعة نرد فقط قطعة واحدة

احتمالية الحصول على رقم 2 يساوي $(\frac{1}{6})$ ناتج واحد من أصل 6 ممكن أن تحدث (1-2-3-4-5-6)

وتساوي $(p=.167)$... واحتمالية عدم الحصول على الرقم 2 يساوي $(\frac{5}{6})$ خمسة احتمالات من أصل

6 $(q=.833)$ لاحظ أن قيمة $p+q=1$

ولاحظ أن $\frac{p}{q}$ تساوي 2. وتعني واحد إلى خمسة (احتمالية الحصول على الرقم 2 مقابل عدم الحصول على الرقم 2)

الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

■ في معادلة خط الانحدار الخطي

■ يكون شكل المعادلة كالتالي:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

■ الإشكالية أن المتغير التابع (المراد التنبؤ به) لا يأخذ سوى إحدى قيمتين (صفر أو واحد)

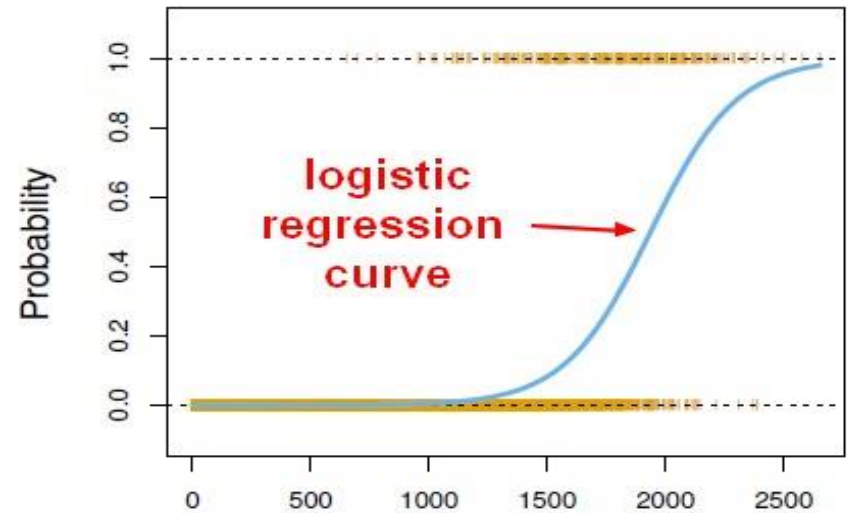
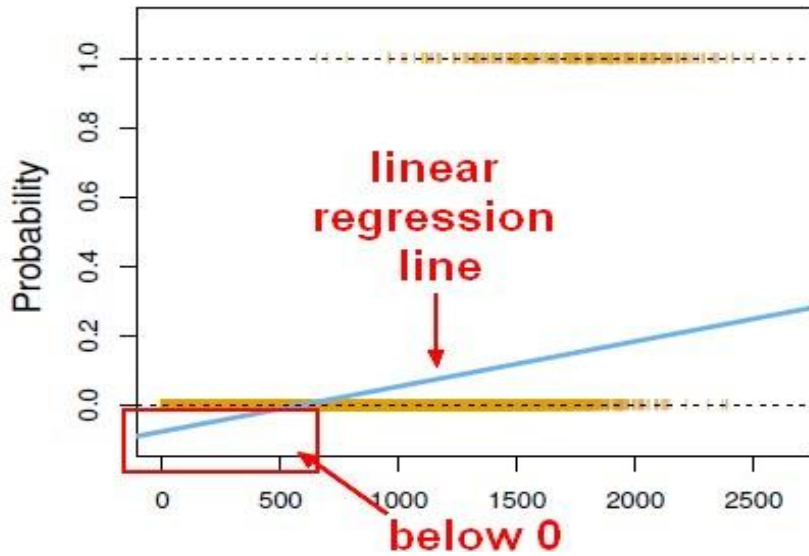
■ أمثلة للمتغير التابع التصنيفي ثنائي المستويات : [مصاب-غير مصاب] [ناجح-راسب]، [مقبول-غير مقبول]، [فوق خط الفقر-تحت خط الفقر].

■ ووفقا لمعادلة خط الانحدار الخطي من الممكن أن تأخذ قيم المتغير التابع قيما أعلى من 1 أو أقل من صفر...

■ ولن يتحقق شرط اعتدالية توزيع الأخطاء (الفرق بين القيم المتوقعة والقيم المشاهدة)

الانحدار اللوغاريتمي Logistic Regression

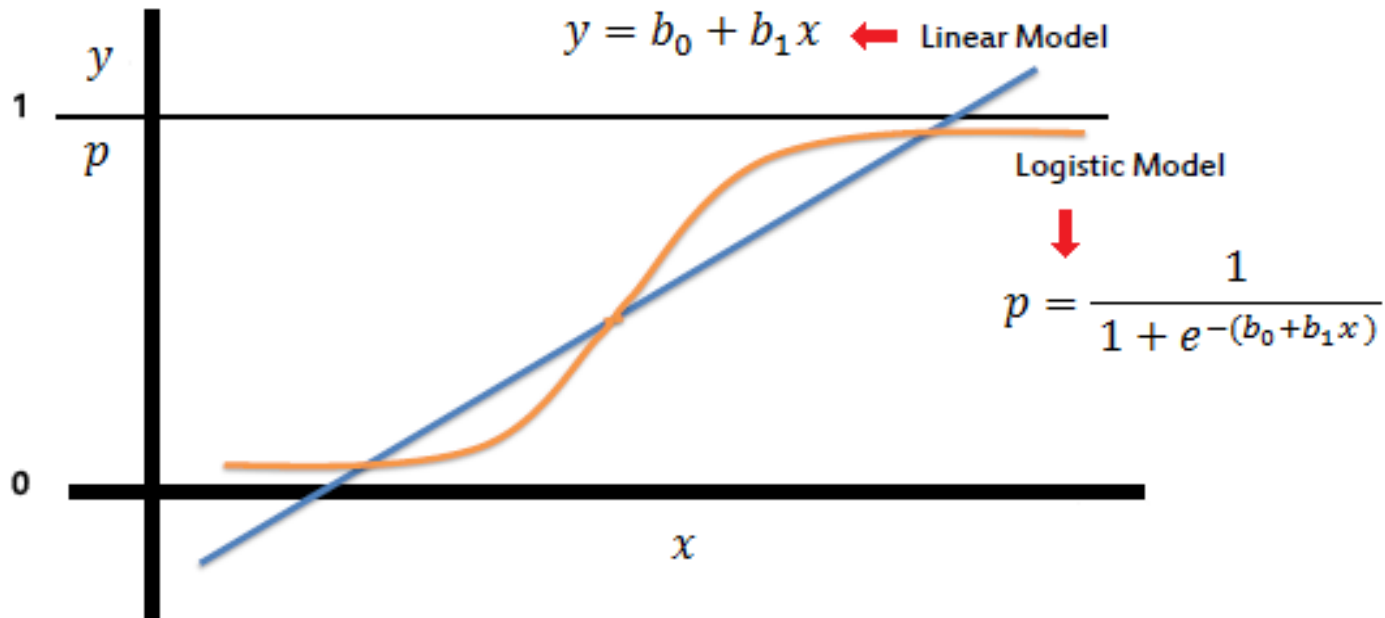
مفاهيم أساسية



الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

■ لذا سنلجأ إلى تحويل رياضية تجعل من قيم المتغير التابع محدودة بين الصفر و الواحد.



الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

بعد إجراء التحويلة الرياضية سنحصل على المعادلة التالية وتسمى لوجت (logit)
ماذا تعني لوجت (logit)؟

تعني اللوغارتم الطبيعي للأرجحية

ستكون معادلة خط الانحدار على النحو التالي: $\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta(\text{Age})$

ولكننا نبحث عن قيمة المتغير التابع (نبحث عن احتمالية قيمة المتغير التابع) بين الصفر و الواحد.
يمكننا ذلك من خلال التحويل الرياضي التالي:

$$\text{Prob}(Y_i = 1) = \frac{e^{\beta'x}}{1 + e^{\beta'x}} \rightarrow \beta_0 + \beta(\text{Age})$$

الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

مثال: التنبؤ بقبول الشخص في الجامعة بمعرفة جنسه (n=20)

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	24.435 ^a	.152	.202

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

Classification Table^a

Observed		Predicted		Percentage Correct
		مقبول	مقبول غير	
Step 1	مقبول غير	7	3	70.0
	مقبول	3	7	70.0
Overall Percentage				70.0

a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a						
(الجنس) 1	1.695	.976	3.015	1	.082	5.444
Constant	-.847	.690	1.508	1	.220	.429

a. Variable(s) entered on step 1: الجنس.

الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	24.435 ^a	.152	.202

a. Estimation terminated at iteration number 3 because parameter estimates changed by less than .001.

مقاييس تحاول مقارنة معامل التحديد
(ينبغي الحذر عند تفسيرها
Cox & Snell
لا يصل للقيمة 1
ولذا ظهر مقياس
Nagelkerke R Square

مثال

Classification Table^a

Observed	Predicted		Percentage Correct	
	مقبول			
	مقبول عبر	مقبول		
مقبول	مقبول عبر	7	3	70.0
	مقبول	3	7	70.0
Overall Percentage				70.0

a. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a الجنس(1)	1.695	.976	3.015	1	.082	5.444
Constant	-.847	.690	1.508	1	.220	.429

a. Variable(s) entered on step 1: الجنس.

الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

مثال:

Variables in the Equation						
	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	(الجنس) 1	1.695	3.015	1	.082	5.444
	Constant	-.847	1.508	1	.220	.429

a. Variable(s) entered on step 1: الجنس.

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta$$

لكن الأكثر فائدة في التفسير هو (**Exp (B)**) ويعني التالي

أرجحية الذكور (حظ الذكور في القبول) 5.44 أضعاف أرجحية الإناث في القبول في الجامعة.

الانحدار اللوغاريتمي Logistic Regression

مفاهيم أساسية

بعد إجراء التحويلة الرياضية سنحصل على المعادلة التالية وتسمى لوجت (logit)
ماذا تعني لوجت (logit)؟

تعني اللوغارتم الطبيعي للأرجحية

ستكون معادلة خط الانحدار على النحو التالي:

$$\ln\left(\frac{P}{1-P}\right) = a + bX$$

$$\frac{P}{1-P} = e^{a+bX}$$

$$P = \frac{e^{a+bX}}{1 + e^{a+bX}}$$

$$\text{Odds} = \frac{\text{probability of the event}}{1 - \text{probability of event}} = \frac{P}{1 - P}$$

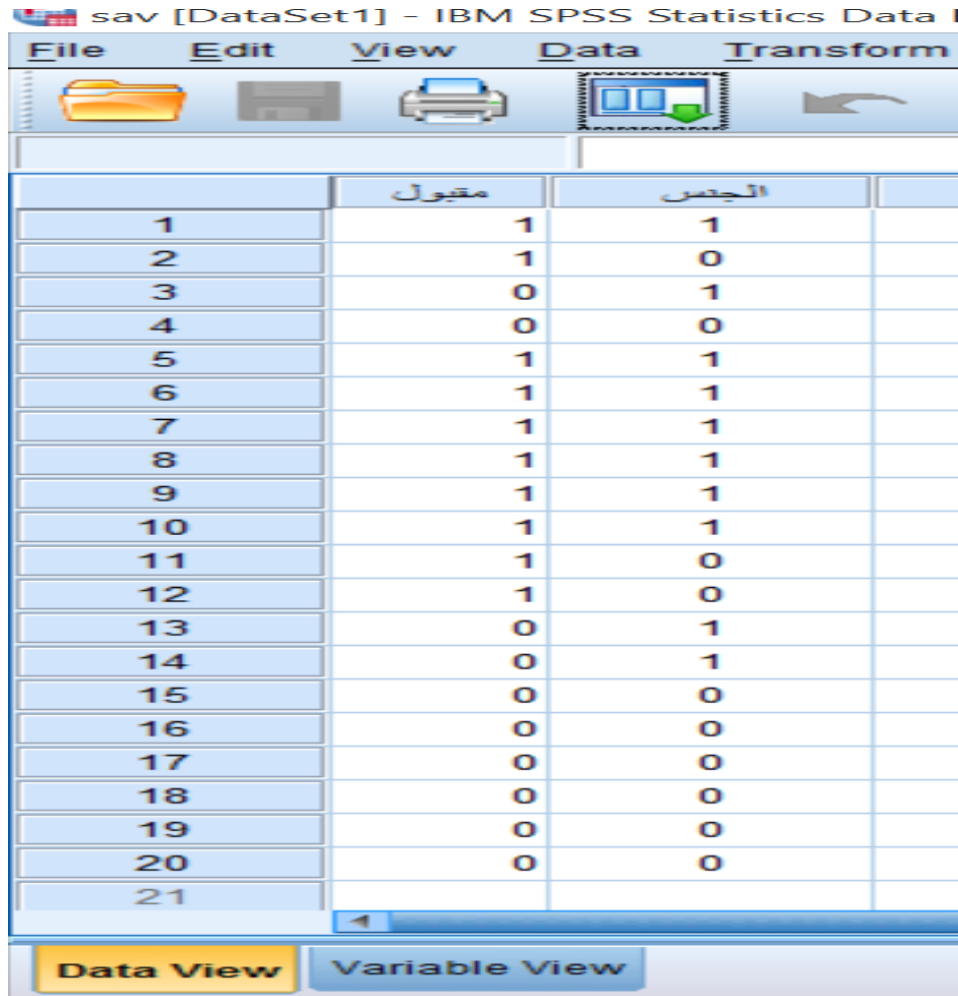
OR

$$P = \frac{\text{Odds}}{1 + \text{Odds}}$$

الانحدار اللوجستي

Logistic Regression

إدخال البيانات



IBM SPSS Statistics Data Editor window showing a dataset with 21 rows and 3 columns. The columns are labeled 'مقبول' (Accepted) and 'الجنس' (Gender). The data is as follows:

	مقبول	الجنس	
1	1	1	
2	1	0	
3	0	1	
4	0	0	
5	1	1	
6	1	1	
7	1	1	
8	1	1	
9	1	1	
10	1	1	
11	1	0	
12	1	0	
13	0	1	
14	0	1	
15	0	0	
16	0	0	
17	0	0	
18	0	0	
19	0	0	
20	0	0	
21			

الانحدار اللوجستي

Logistic Regression

إدخال البيانات

القبول للذكور وال (1) Edit

The screenshot shows the SPSS software interface. The 'Analyze' menu is open, and the 'Regression' option is selected. The 'Binary Logistic...' option is highlighted in the sub-menu. A 'Shift' key icon is visible next to the 'Binary Logistic...' option.

	var	var	var
.42857			
.33333			
.33333			
.42857			

الانحدار اللوجستي

Logistic Regression

إدخال البيانات

Logistic Regression

Dependent: مقبول ①

Block 1 of 1

Previous Next

Covariates: الجنس(Cat) ②

Method: Enter

Selection Variable: Rule...

OK Paste Reset Cancel Help

Categorical ③

Save... Options... Style... Bootstrap...

الجنس

Predicted probability...

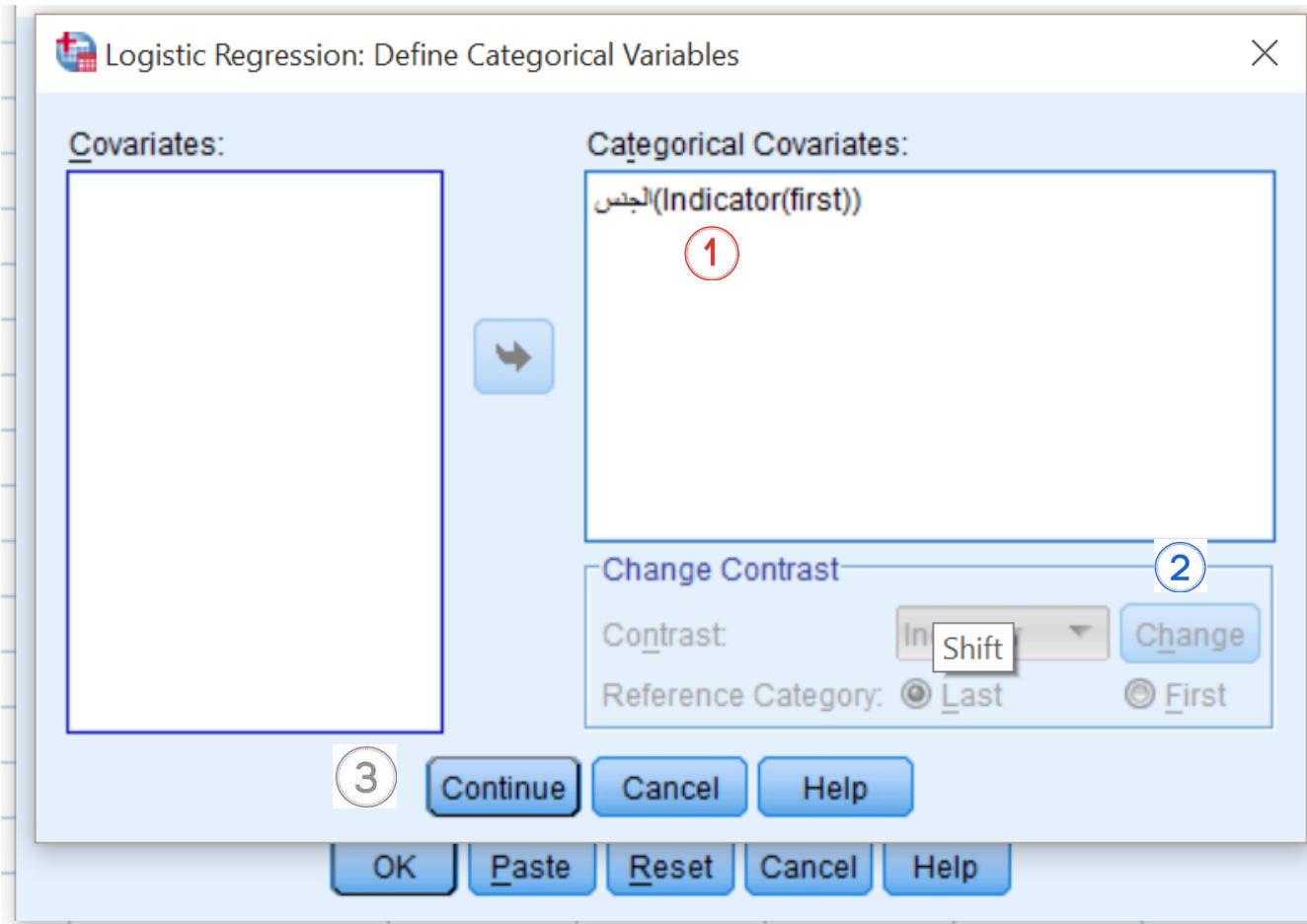
Predicted group [PG...]

Logit residual [LRE...]

الانحدار اللوجستي

Logistic Regression

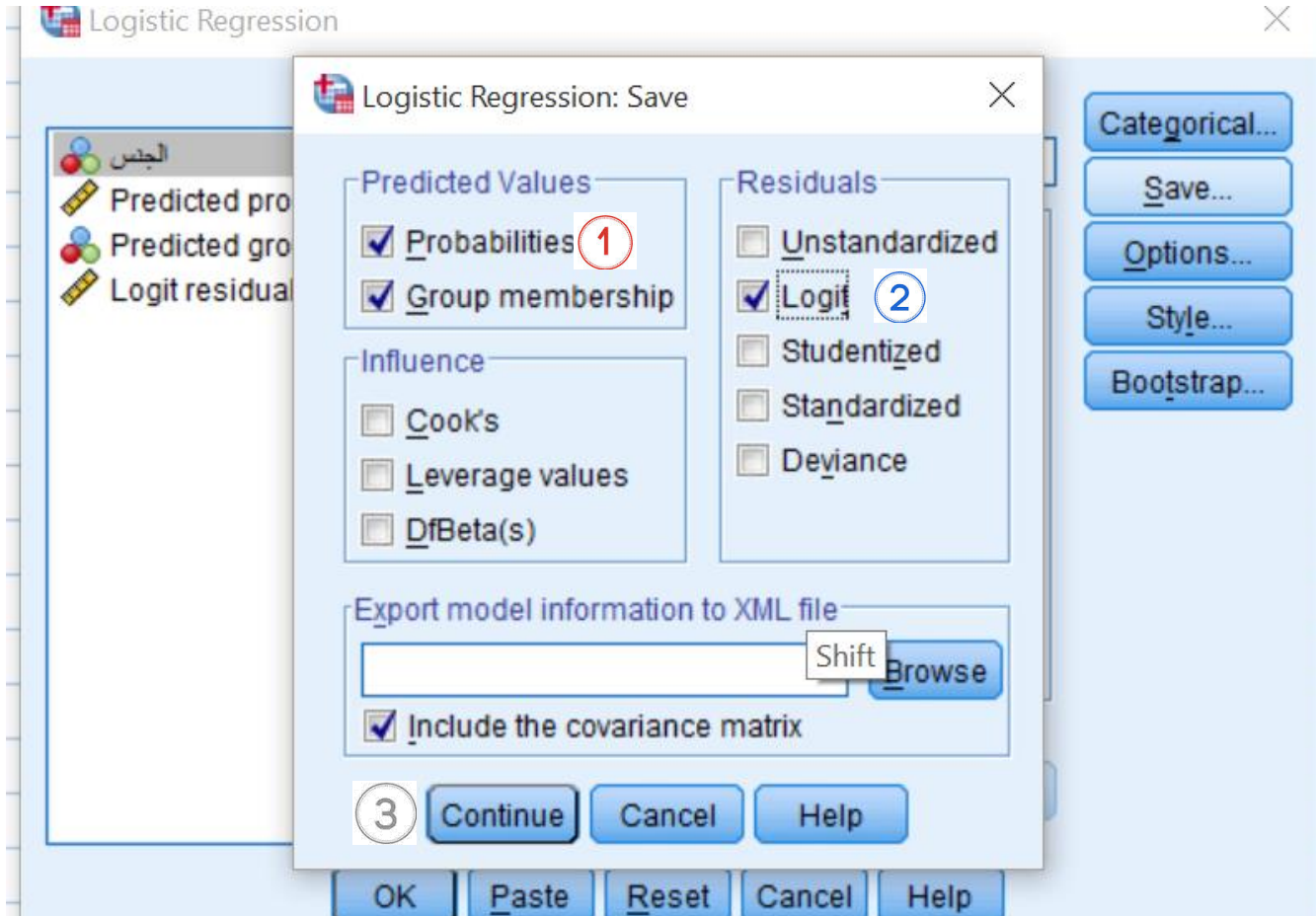
إدخال البيانات



الانحدار اللوجستي

Logistic Regression

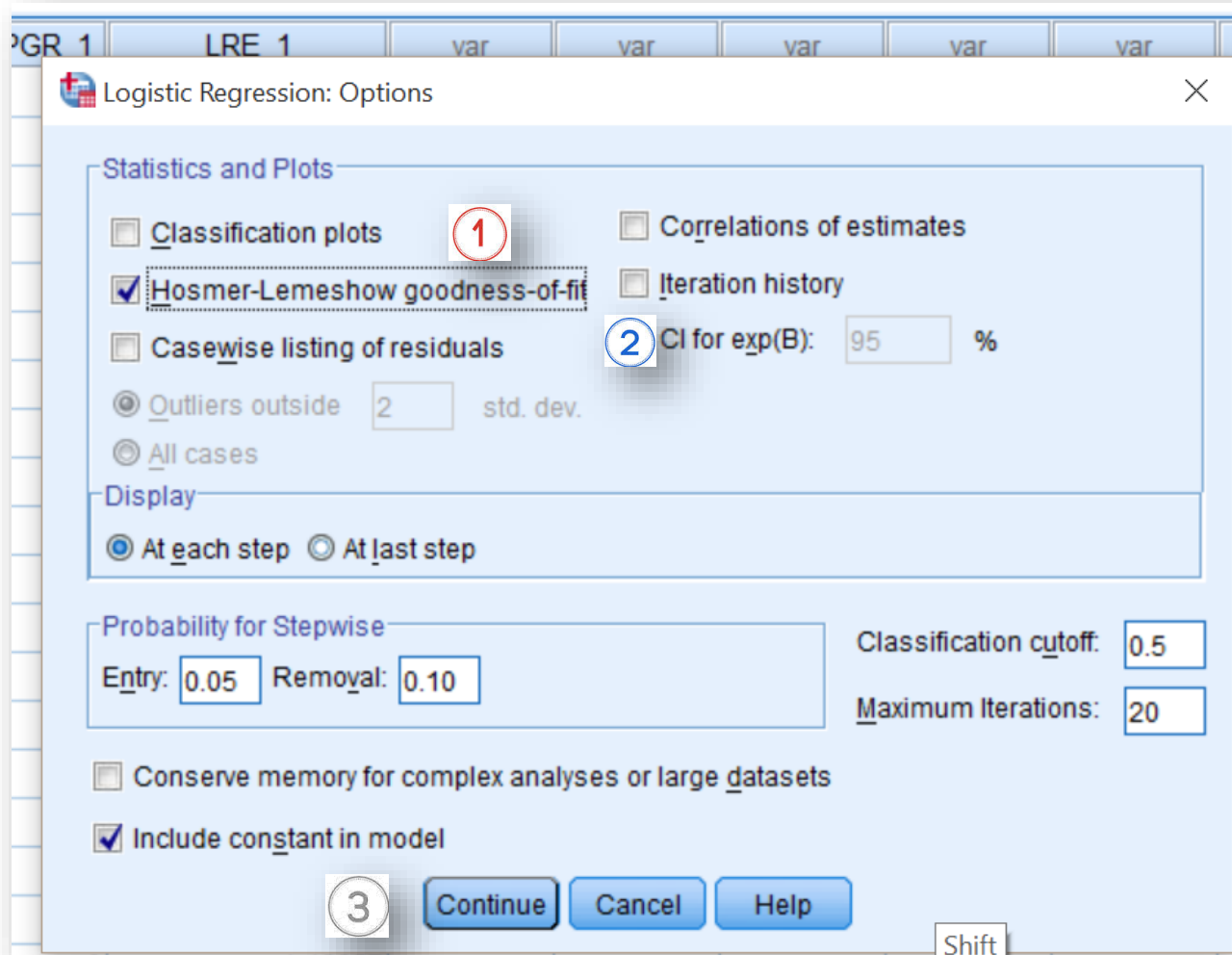
إدخال البيانات 



الانحدار اللوجستي

Logistic Regression

إدخال البيانات 



Logistic Regression: Options

Statistics and Plots

- Classification plots 1
- Hosmer-Lemeshow goodness-of-fit
- Casewise listing of residuals
- Correlations of estimates
- Iteration history
- CI for exp(B): % 2
- Outliers outside std. dev.
- All cases

Display

- At each step At last step

Probability for Stepwise

Entry: Removal:

Classification cutoff:

Maximum iterations:

Conserve memory for complex analyses or large datasets

Include constant in model

3

Shift

الانحدار اللوجستي

Logistic Regression

إدخال البيانات

Logistic Regression

Dependent: مقبول

Block 1 of 1

Previous Next

Covariates: الجنس (Cat)

Method: Enter

Selection Variable:

OK Paste Reset Cancel Help

1