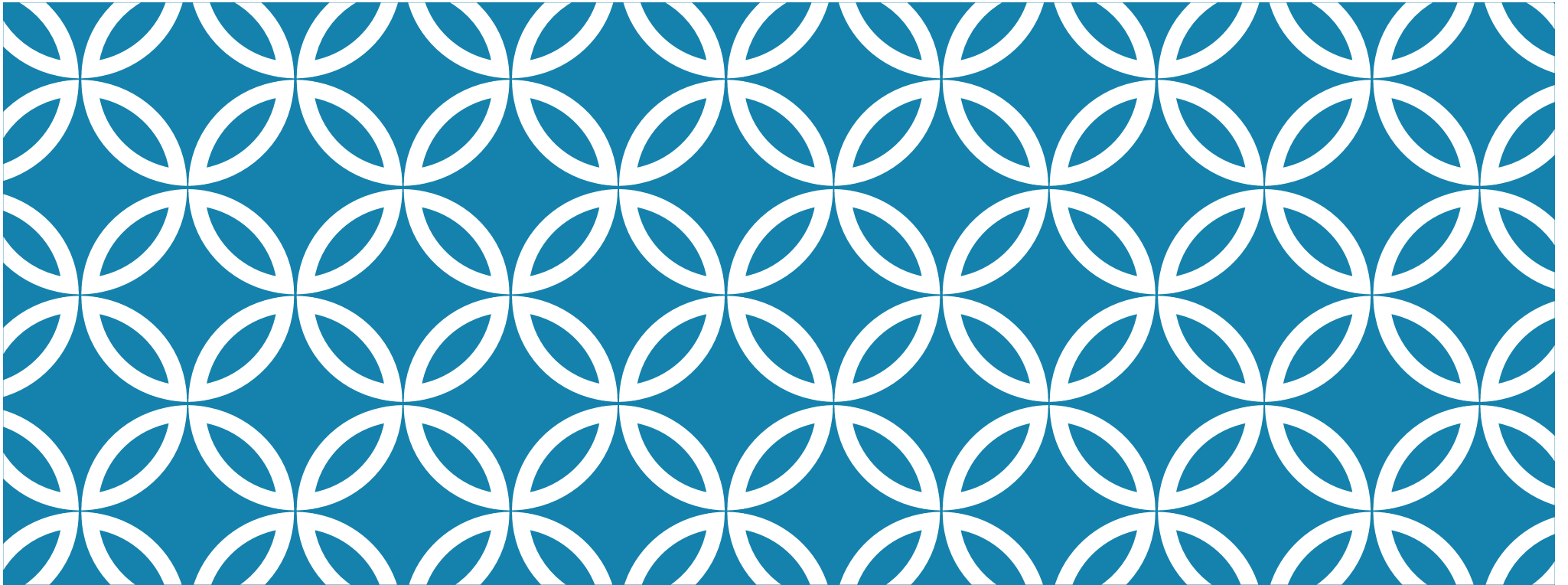




LOGISTIC MODEL CLASS EXERCISE

MB628

ตัวอย่างการวิเคราะห์การถดถอยโลจิสติก

ผู้วิจัยสนใจทำนายการเป็นโรคมะเร็งต่อมน้ำเหลืองโดยมีที่มาของปัญหาการวิจัยว่า การเป็นมะเร็งนั้นมีทฤษฎีว่าอาจมาจากหรือขึ้นอยู่กับกลไกของเซลล์มะเร็ง ซึ่งมักเกิดในต่อมน้ำเหลืองก่อนการตรวจสอบมะเร็งในระยะแรก จึงต้องอาศัยการผ่าตัดต่อมน้ำเหลืองออกเพื่อสำรวจว่าเซลล์มะเร็งยังอยู่หรือไม่แล้วจึงทำการรักษาที่จำเป็นต่อไป ถ้าผู้วิจัยทำนายการมีเซลล์มะเร็งในต่อมน้ำเหลืองได้โดยอาศัยข้อมูลการวินิจฉัยที่มีอยู่ก็จะไม่จำเป็นต้องทำการผ่าตัดต่อมน้ำเหลือง ซึ่งจะทำให้ลดค่าใช้จ่ายต่าง ๆ ออกไปได้มาก ผู้วิจัยจึงเก็บรวบรวมข้อมูลจากผู้เข้ารักษามะเร็งในต่อมน้ำเหลืองจำนวน 160 คน โดยเก็บข้อมูลตามตัวแปรทำนายดังนี้

❖ Age อายุ

❖ X-ray ผล

❖ Grade ระดับความผิดปกติของเซลล์

❖ Stage ระยะโรค

❖ Acid (serum acid phosphatase) สารนี้จะมีค่าสูง เมื่อมีเซลล์ผิดปกติ

ตัวแปรเกณฑ์ในการวิจัยนี้คือ อาการของโรคมะเร็งต่อมน้ำเหลือง แบ่งเป็น 2 กลุ่มดังนี้

0 : ไม่เป็นโรคมะเร็งในต่อมน้ำเหลือง

1 : เป็นโรคมะเร็งในต่อมน้ำเหลือง

LOGITYUTH1.SAV

1. เลือก analyze
2. เลือก regression
3. คลิกที่ binary logistic

logityuth1.sav - SPSS Data Editor																	
File Edit View Data Transform Analyze Graphs Utilities Window Help																	
1 : age 45																	
	age	xray	grade	stage	acid	CHD	var	var	var	var	var	var	var	var	var	var	var
1	45.00	.00	2.00	3.00	45.00	.00											
2	52.00	1.00	2.00	3.00	48.00	1.00											
3	50.00	1.00	3.00	4.00	67.00	.00											
4	60.00	.00	2.00	2.00	50.00	.00											
5	68.00	1.00	2.00	2.00	50.00	1.00											
6	45.00	1.00	1.00	2.00	45.00	1.00											
7	51.00	.00	2.00	2.00	50.00	1.00											
8	56.00	1.00	2.00	3.00	54.00	.00											
9	45.00	1.00	3.00	4.00	64.00	1.00											
10	40.00	.00	4.00	5.00	76.00	1.00											
11	30.00	1.00	1.00	1.00	34.00	1.00											
12	68.00	1.00	2.00	3.00	60.00	.00											
13	65.00	.00	4.00	4.00	78.00	1.00											
14	69.00	1.00	3.00	4.00	67.00	1.00											
15	61.00	.00	4.00	5.00	75.00	1.00											
16	65.00	1.00	5.00	5.00	86.00	.00											
17	61.00	1.00	2.00	2.00	45.00	1.00											
18	45.00	.00	3.00	3.00	60.00	1.00											
19	41.00	1.00	4.00	4.00	70.00	.00											
20	65.00	1.00	5.00	5.00	80.00	1.00											
21	56.00	.00	4.00	5.00	71.00	1.00											
22	45.00	1.00	1.00	2.00	34.00	.00											
23	54.00	.00	2.00	3.00	43.00	.00											
24	65.00	1.00	2.00	3.00	54.00	1.00											
25	67.00	1.00	2.00	3.00	46.00	1.00											
26	62.00	.00	3.00	4.00	50.00	.00											
27	60.00	1.00	1.00	2.00	34.00	1.00											
28	53.00	.00	2.00	2.00	42.00	.00											
29	56.00	.00	2.00	2.00	40.00	1.00											
30	55.00	1.00	2.00	2.00	45.00	.00											
31	45.00	.00	2.00	3.00	45.00	.00											
32	65.00	1.00	5.00	5.00	80.00	1.00											

Logistic Regression

age

xray

grade

stage

acid

Dependent:

CHD

Block 1 of 1

Previous

Next

Covariates:

age

xray

grade

stage

acid

Method:

Enter

Selection Variable:

Rule...

Categorical...

Save...

Options...

OK

Paste

Reset

Cancel

Help

logityuth1.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : age 45

	age	xray	grade	stage	acid	CHD	var	var	var	var	var	var	var	var	var	var
1	45.00	.00	2.00	3.00	45.00	.00										
2	52.00	1.00	2.00	3.00	48.00	1.00										
3	50.00	1.00	3.00	4.00	67.00	.00										
4	60.00	.00	2.00	2.00	50.00	.00										
5	68.00	1.00	2.00	2.00	50.00	1.00										
6	45.00	1.00	1.00	2.00	45.00	1.00										
7	51.00	.00	2.00	2.00	50.00	1.00										
8	56.00	1.00	2.00	3.00	54.00	.00										
9	45.00	1.00	3.00	4.00	64.00	1.00										
10	40.00	.00	4.00	5.00	76.00	1.00										
11	30.00	1.00	1.00	1.00	34.00	1.00										
12	68.00	1.00	2.00	3.00	60.00	.00										
13	65.00	.00	4.00	4.00	78.00	1.00										
14	69.00	1.00	3.00	4.00	67.00	1.00										
15	61.00	.00	4.00	5.00	75.00	1.00										
16	65.00	1.00	5.00	5.00	86.00	.00										
17	61.00	1.00	2.00	2.00	45.00	1.00										
18	45.00	.00	3.00	3.00	60.00	1.00										
19	41.00	1.00	4.00	4.00	70.00	.00										
20	65.00	1.00	5.00	5.00	80.00	1.00										
21	56.00	.00	4.00	5.00	71.00	1.00										
22	45.00	1.00	1.00	2.00	34.00	.00										
23	54.00	.00	2.00	3.00	43.00	.00										
24	65.00	1.00	2.00	3.00	54.00	1.00										
25	67.00	1.00	2.00	3.00	46.00	1.00										
26	62.00	.00	3.00	4.00	50.00	.00										
27	60.00	1.00	1.00	2.00	34.00	1.00										
28	53.00	.00	2.00	2.00	42.00	.00										
29	56.00	.00	2.00	2.00	40.00	1.00										
30	55.00	1.00	2.00	2.00	45.00	.00										
31	45.00	.00	2.00	3.00	45.00	.00										
32	65.00	1.00	5.00	5.00	80.00	1.00										

Logistic Regression

Dependent: CHD

Block 1 of 1

Covariates: age, xray, grade, stage, acid

Method: Enter

Selection Variable:

Logistic Regression: Save

Predicted Values: ☒ Probabilities, ☒ Group membership

Influence: ☒ Cook's, ☐ Leverage values, ☐ DfBeta(s)

Residuals: ☐ Unstandardized, ☐ Logit, ☐ Studentized, ☒ Standardized, ☒ Deviance

Export model information to XML file: Browse

☒ Include the covariance matrix

logityuth1.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : age 45

	age	xray	grade	stage	acid	CHD	var	var	var	var	var	var	var	var	var	var
1	45.00	.00	2.00	3.00	45.00	.00										
2	52.00	1.00	2.00	3.00	48.00	1.00										
3	50.00	1.00	3.00	4.00	67.00	.00										
4	60.00	.00	2.00	2.00	50.00	.00										
5	68.00	1.00	2.00	2.00	50.00	1.00										
6	45.00	1.00	1.00	2.00	45.00	1.00										
7	51.00	.00	2.00	2.00	50.00	1.00										
8	56.00	1.00	2.00	3.00	54.00	.00										
9	45.00	1.00	3.00	4.00	64.00	1.00										
10	40.00	.00	4.00	5.00	76.00	1.00										
11	30.00	1.00	1.00	1.00	34.00	1.00										
12	68.00	1.00	2.00	3.00	60.00	.00										
13	65.00	.00	4.00	4.00	78.00	1.00										
14	69.00	1.00	3.00	4.00	67.00	1.00										
15	61.00	.00	4.00	5.00	75.00	1.00										
16	65.00	1.00	5.00	5.00	86.00	.00										
17	61.00	1.00	2.00	2.00	45.00	1.00										
18	45.00	.00	3.00	3.00	60.00	1.00										
19	41.00	1.00	4.00	4.00	70.00	.00										
20	65.00	1.00	5.00	5.00	80.00	1.00										
21	56.00	.00	4.00	5.00	71.00	1.00										
22	45.00	1.00	1.00	2.00	34.00	.00										
23	54.00	.00	2.00	3.00	43.00	.00										
24	65.00	1.00	2.00	3.00	54.00	1.00										
25	67.00	1.00	2.00	3.00	46.00	1.00										
26	62.00	.00	3.00	4.00	50.00	.00										
27	60.00	1.00	1.00	2.00	34.00	1.00										
28	53.00	.00	2.00	2.00	42.00	.00										
29	56.00	.00	2.00	2.00	40.00	1.00										
30	55.00	1.00	2.00	2.00	45.00	.00										
31	45.00	.00	2.00	3.00	45.00	.00										
32	65.00	1.00	5.00	5.00	80.00	1.00										

Logistic Regression

Dependent: CHD

Block 1 of 1

Covariates: age, xray, grade, stage, acid

Method: Enter

Selection Variable:

Logistic Regression: Options

Statistics and Plots

- ☒ Classification plots
- ☒ Hosmer-Lemeshow goodness-of-fit
- ☐ Casewise listing of residuals
- ☐ Outliers outside 2 std. dev.
- ☐ All cases
- ☒ Correlations of estimates
- ☒ Iteration history
- ☒ CI for exp(B): 95 %

Display

- ☒ At each step
- ☐ At last step

Probability for Stepwise

Entry: .05 Removal: .10

Classification cutoff: .5

Maximum iterations: 20

ตารางแสดงจำนวนของตัวอย่างและเปอร์เซ็นต์

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	160	100.0
	Missing Cases	0	.0
	Total	160	100.0
Unselected Cases		0	.0
Total		160	100.0

Case ที่นำมาวิเคราะห์

Missing cases ไม่มี

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

ตารางแสดงค่าตัวแปรเกณฑ์

Dependent Variable Encoding

Original Value	Internal Value
.00	0
1.00	1

BLOCK 0: BEGINNING BLOCK

Iteration History^{a,b,c} ← การทำงานหลายๆ รอบ

Iteration		-2 Log likelihood	Coefficients
			Constant
Step 1		220.906	.150
0	2	220.906	.150

-2LL เริ่มต้น = 220.906

- a. Constant is included in the model.
- b. Initial -2 Log Likelihood: 220.906
- c. Estimation terminated at iteration number 2 because parameter estimates changed by less than .001.

Step 0 จำนวน 2 รอบ

ผลทำนายที่ยังไม่นำตัวแปรเข้าสมการ

มีเฉพาะค่าคงที่มีความถูกต้อง 53.8%

Classification Table^b

Observed			Predicted		Percentage Correct
			CHD		
			.00	1.00	
Step 0	CHD	.00	0	74	.0
		1.00	0	86	100.0
Overall Percentage					53.8

- a. Constant is included in the model.
- b. The cut value is .500

ผลลัพธ์ของตัวแปรที่อยู่ในสมการของ STEP 0

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	.150	.159	.898	1	.343	1.162

ค่าคงที่ไม่มีผลต่อการทำนาย

ค่าตัวแปรทำนายและยังไม่ได้นำเข้าสมการ

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	age	3.892	1	.049
		xray	.507	1	.476
		grade	13.804	1	.000
		stage	9.006	1	.003
		acid	17.229	1	.000
Overall Statistics			23.034	5	.000

BLOCK 1: METHOD=ENTER

Iteration History^{a,b,c,d}

		-2 Log likelihood	Coefficients					
Iteration			Constant	age	xray	grade	stage	acid
Step 1	1	196.734	-.437	-.035	.178	-.120	-.306	.069
	2	196.382	-.346	-.042	.210	-.106	-.346	.076
	3	196.382	-.337	-.042	.211	-.104	-.347	.076
	4	196.382	-.337	-.042	.211	-.104	-.347	.076

a. Method: Enter

b. Constant is included in the model.

c. Initial -2 Log Likelihood: 220.906

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

ตัวแปรทำนายร่วมกัน

ทำนายโอกาสเกิด

โรคมะเร็งได้

H_0 : โอกาสที่จะเป็นโรคมะเร็งต่อมน้ำเหลืองไม่ขึ้นอยู่กับตัวแปรทั้ง 5 ตัวแปร

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	24.525	5	.000
	Block	24.525	5	.000
	Model	24.525	5	.000

MODEL SUMMARY

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	196.382 ^a	.142	.190

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

สถิติทดสอบความสอดคล้อง model คิดเป็น 14.2%

ค่า -2LL ที่มีการนำตัวแปรอิสระ 5 ตัวเข้าสมการแล้ว

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	6.339	8	.609

ยอมรับสมมติฐานหลัก คือ Model มีความเหมาะสม

CLASSIFICATION TABLE

Classification Table

Observed			Predicted		Percentage Correct
			CHD		
			.00	1.00	
Step 1	CHD	.00	42	32	56.8
		1.00	22	64	74.4
Overall Percentage					66.3

a. The cut value is .500

VARIABLES IN THE EQUATION

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
Step 1	age	-.042	.021	3.951	1	.047	.959	.920	.999
	xray	.211	.347	.371	1	.543	1.235	.626	2.440
	grade	-.104	.460	.051	1	.821	.901	.365	2.222
	stage	-.347	.279	1.552	1	.213	.707	.409	1.220
	acid	.076	.036	4.586	1	.032	1.079	1.007	1.158
	Constant	-.337	1.437	.055	1	.815	.714		

a. Variable(s) entered on step 1: age, xray, grade, stage, acid.

มีนัยสำคัญแสดงว่าตัวนี้มีผลต่อการเป็นโรคมะเร็ง

- ❖ ค่าสถิติ Wald statistics พบว่า ตัวแปรทำนาย 2 ตัว ได้แก่ age และ acid ที่มีผลต่อการเป็นโรคมะเร็งต่อมน้ำเหลืองอย่างมีนัยสำคัญทางสถิติที่ ระดับ 0.05
- ❖ ค่า $\exp(B)$ น้อยกว่า 1 จะทำให้ลดโอกาสการเกิดเหตุการณ์ที่สนใจ
- ❖ ค่า $\exp(B)$ มากกว่า 1 จะทำให้โอกาสการเกิดเหตุการณ์ที่สนใจเพิ่มขึ้น

CORRELATION MATRIX

Correlation Matrix

		Constant	age	xray	grade	stage	acid
Step 1	Constant	1.000	-.818	-.069	.346	.090	-.464
	age	-.818	1.000	-.027	-.036	.013	-.004
	xray	-.069	-.027	1.000	.128	-.040	-.090
	grade	.346	-.036	.128	1.000	-.194	-.777
	stage	.090	.013	-.040	-.194	1.000	-.383
	acid	-.464	-.004	-.090	-.777	-.383	1.000