

Probit Regression Workshop #1-GENLIN syntax command

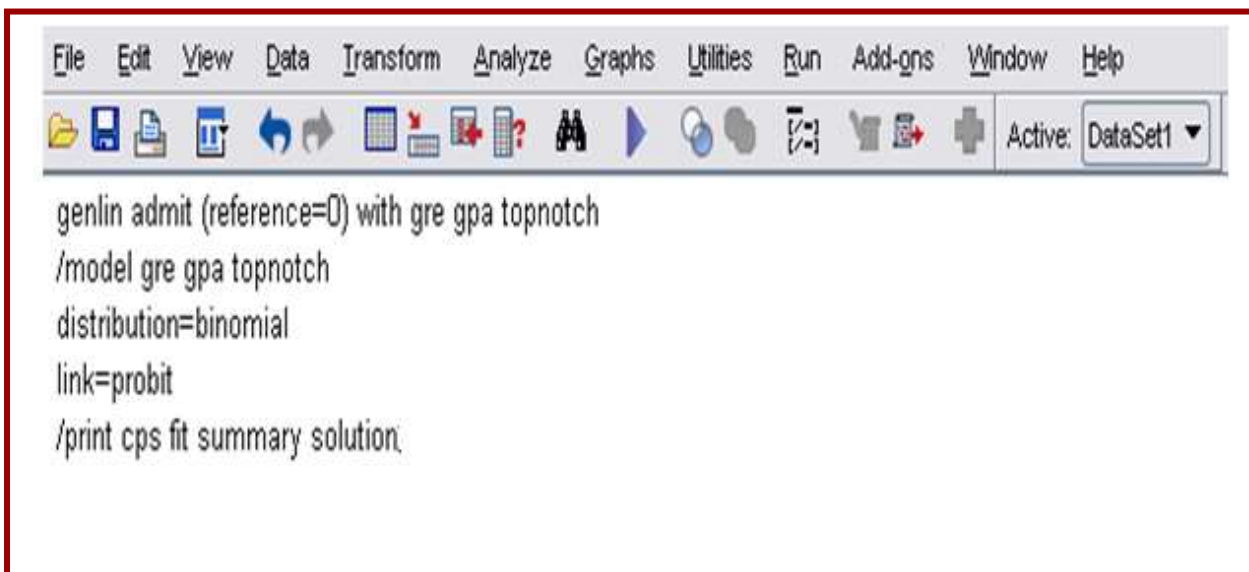
ตารางด้านล่างแสดงข้อมูลที่ประกอบด้วยตัวแปรอิสระ 3 ตัวได้แก่ คะแนนGRE (Graduate Record examination test) เกรดเฉลี่ยสะสม (GPA) และชื่อเสียงของโรงเรียนที่นักศึกษาจบการศึกษา ส่วนตัวแปรตามได้แก่การได้รับการตอบรับเข้าเรียนในระดับปริญญาตรี

	admit	gre	topnotch	gpa
1	0	380.00	0	3.61
2	1	660.00	1	3.67
3	1	800.00	1	4.00
4	1	640.00	0	3.19
5	0	520.00	0	2.93
6	1	760.00	0	3.00
7	1	560.00	0	2.98
8	0	400.00	0	3.08
9	1	540.00	0	3.39
10	0	700.00	1	3.92
11	0	800.00	1	4.00
12	0	440.00	0	3.22
13	1	760.00	0	4.00
14	0	700.00	0	3.08
15	1	700.00	1	4.00
16	0	480.00	0	3.44
17	0	780.00	0	3.87
18	0	360.00	0	2.56
19	0	800.00	0	3.75
20	1	540.00	0	3.81
21	0	500.00	0	3.17
22	1	660.00	1	3.63
23	0	600.00	0	2.82
24	0	680.00	0	3.19
25	1	760.00	0	3.35
26	1	800.00	0	3.66
27	1	620.00	1	3.61
28	1	520.00	0	3.74
29	1	780.00	0	3.22
30	0	520.00	0	3.29
31	0	540.00	0	3.78
32	0	760.00	0	3.35
33	0	600.00	0	3.40
34	1	800.00	1	4.00
35	0	360.00	0	3.14
36	0	400.00	0	3.05
37	0	580.00	0	3.25
38	0	520.00	0	2.90
39	1	500.00	0	3.13

- คลิก File>New>Syntax ดังภาพต่อไปนี้



จากนั้นพิมพ์ GENLIN syntax ลงใน syntax editor ดังภาพต่อไปนี้



•คลิก Run>All จะปรากฏ computer output ตามลำดับดังนี้

1. Case processing summary

	N	Percent
Included	400	100.0%
Excluded	0	.0%
Total	400	100.0%

Case processing summary แสดงให้เห็นขนาดของกลุ่มตัวอย่าง

2. Goodness of fit table

	Value	df	Value/df
Deviance	440.291	370	1.190
Scaled Deviance	440.291	370	
Pearson Chi-Square	368.411	370	.996
Scaled Pearson Chi-Square	368.411	370	
Log Likelihood ^a	-238.943		
Akaike's Information Criterion (AIC)	485.887		
Finite Sample Corrected AIC (AICC)	485.988		
Bayesian Information Criterion (BIC)	501.853		
Consistent AIC (CAIC)	505.853		

Dependent Variable: admit
Model: (Intercept), gre, gpa, topnotch

a. The full log likelihood function is displayed and used in computing information criteria.

b. Information criteria are in small-is-better form.

Dispersion parameter วัดจาก Deviance / d.f. มีค่าเข้าใกล้ 1 บ่งบอกความ fit ของ model ที่ใช้

Likelihood Ratio Chi-Square	df	Sig.
22.090	3	.000

Dependent Variable: admit
Model: (Intercept), gre, gpa, topnotch

a. Compares the fitted model against the intercept-only model.

ผลจากการใช้ full model (intercept+ตัวแปร) มีผลทำให้ Deviance หรือความผันผวนของตัวแปรตามที่ไม่สามารถอธิบายได้ลดลงจาก Deviance ของตัวแปรตามที่ใช้ intercept model อย่างมีนัยสำคัญ

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	18.669	1	.000
gre	5.706	1	.017
gpa	4.312	1	.038
topnotch	2.311	1	.128

Dependent Variable: admit
Model: (Intercept), gre, gpa, topnotch

Test of Model Effects แสดงให้เห็นว่าตัวแปรที่มีค่าต่อเนื่อง (gre & gpa) มีผลต่อตัวแปรตามอย่างมีนัยสำคัญ ในขณะที่ตัวแปรตามที่เป็นนามบัญญัติ (topnotch) ไม่มีผลอย่างมีนัยสำคัญ

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-2.798	.6475	-4.067	-1.529	18.669	1	.000
gre	.002	.0006	.000	.003	5.706	1	.017
gpa	.401	.1931	.023	.779	4.312	1	.038
topnotch (Scale)	.273 ^a	.1796	-.079	.625	2.311	1	.128

Dependent Variable: admit
Model: (Intercept), gre, gpa, topnotch

a. Fixed at the displayed value.

สัมประสิทธิ์หน้าตัวแปรที่มีค่าเป็นบวกมีความหมายว่าการเพิ่มขึ้นของค่าตัวแปรอิสระจะมีผลทำให้ความน่าจะเป็นเพิ่มขึ้น ในทางกลับกันสัมประสิทธิ์หน้าตัวแปรที่มีค่าเป็นลบมีความหมายว่าการเพิ่มขึ้นของค่าตัวแปรอิสระจะมีผลทำให้ความน่าจะเป็นลดลง ในที่นี้ การเพิ่มขึ้นของ gre ทำให้ความน่าจะเป็นที่จะได้รับการตอบรับเข้าศึกษามีสูงขึ้น การได้เรียนอยู่ในสถาบันที่มีชื่อเสียง (topnotch=1) จะทำให้ความน่าจะเป็นที่จะได้รับการตอบรับเข้าศึกษามีสูงขึ้น และการเพิ่มขึ้นของ gpa ทำให้ความน่าจะเป็นที่จะได้รับการตอบรับเข้าศึกษามีสูงขึ้น อย่างไรก็ตามเนื่องจากการตีความของ probit regression ไม่ได้มีความหมายตรงไปตรงมาเหมือนกับการตีความใน logistic regression หรือ ordinary linear regression ทำให้เราไม่สามารถระบุตัวเลขความน่าจะเป็นที่เพิ่มขึ้นได้แน่นอน

หากใน syntax command เราเพิ่มเติมคำสั่ง save predval ดังแสดงได้ดังภาพในหน้าถัดไป

```

File Edit View Data Transform Analyze Graphs Utilities Run Add-ons Window Help
genlin admit(reference=0) with gre gpa topnotch
/model gre gpa topnotch
distribution=binomial
link=probit
/print cps fit summary solution
/save predval.

```

ในตาราง data editor จะปรากฏค่าของตัวแปรตามที่ได้จากการพยากรณ์ในคอลัมน์สุดท้ายของตาราง

ค่าของตัวแปรตามที่ได้จากการพยากรณ์

	admit	gre	topnotch	gpa	PredictedValue
1	0	380.00	0	3.61	0
2	1	660.00	1	3.67	0
3	1	800.00	1	4.00	1
4	1	640.00	0	3.19	0
5	0	520.00	0	2.93	0
6	1	760.00	0	3.00	0
7	1	560.00	0	2.98	0
8	0	400.00	0	3.08	0
9	1	540.00	0	3.39	0
10	0	700.00	1	3.92	1
11	0	800.00	1	4.00	1
12	0	440.00	0	3.22	0
13	1	760.00	0	4.00	0
14	0	700.00	0	3.08	0
15	1	700.00	1	4.00	1
16	0	480.00	0	3.44	0
17	0	780.00	0	3.87	0
18	0	360.00	0	2.56	0
19	0	800.00	0	3.75	0
20	1	540.00	0	3.81	0
21	0	500.00	0	3.17	0
22	1	660.00	1	3.63	0
23	0	600.00	0	2.82	0
24	0	680.00	0	3.19	0
25	1	760.00	0	3.35	0
26	1	800.00	0	3.66	0
27	1	620.00	1	3.61	0
28	1	520.00	0	3.74	0
29	1	780.00	0	3.22	0
30	0	520.00	0	3.29	0
31	0	540.00	0	3.78	0
32	0	760.00	0	3.35	0
33	0	600.00	0	3.40	0
34	1	800.00	1	4.00	1
35	0	360.00	0	3.14	0
36	0	400.00	0	3.05	0
37	0	580.00	0	3.25	0

จากการทำตารางไขว้ (cross tabulation) สามารถสร้าง classification table ได้ดังนี้

admit * Predicted Category Value Crosstabulation

Count		Predicted Category Value		Total
		0	1	
admit	0	260	13	273
	1	115	12	127
Total		375	25	400

ความถูกต้องของการพยากรณ์= 68%
 $[(260+12)/400]*100$

Contribution this issue: ดร. ดนัย ปัตตพงศ์

อยากเรียนรู้การนำสถิติข้างต้นนี้ไปใช้ในการวิจัยระดับสารนิพนธ์ (independent study) วิทยานิพนธ์ (thesis) ดุษฎีนิพนธ์(dissertation) ปรึกษาได้ที่ dpattaphongse@gmail.com

- * ผู้แต่ง MBA's Made Easy (160+ issues) เอกสารวิชาการด้านศาสตร์การบริหารธุรกิจที่ช่วยให้ธุรกิจสามารถยืนหยัดและอยู่รอดได้ในภาวะที่โลกเปลี่ยนแปลงอยู่ตลอดเวลา
- * ผู้พัฒนา FINALYSIS... a dedicated software สำหรับให้บริการนักธุรกิจที่ต้องการวิเคราะห์ความเป็นไปได้ทางการเงินของโครงการพัฒนาอสังหาริมทรัพย์ (บ้านจัดสรร/จัดสรรที่ดินเพื่อการอุตสาหกรรม/อาคารชุด/อาคารสำนักงานให้เช่า) โรงแรม โรงพยาบาลเอกชน ห้างสรรพสินค้า โรงงานน้ำตาล โรงงานกระดาษ โรงไฟฟ้าชีวมวล ฯลฯ ได้เห็นตัวเลขก่อนโครงการเกิด หลีกเลียงความผิดพลาดเป็นร้อยเป็นพันล้านหากเกิดการลงทุนจริง(กำหนด DEBUT 1 เมษายน 2569)
- * ผู้แต่งหนังสือ”การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน”ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม
- * Co-developer ซอฟต์แวร์ en@gex@cel[®] สำหรับใช้ทดสอบ/เรียนรู้ศัพท์(ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 90 บทครอบคลุมศัพท์ระดับ SAT/IELTS/TOEFL กว่า 12,000 คำ) และไวยากรณ์อังกฤษ (ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 160 บทหรือกว่า 10,000 ข้อครอบคลุมเนื้อหาาระดับอุดมศึกษาและTOEFL) มาพร้อมกับไฟล์เสียง/ไฟล์ข้อมูล/ฯลฯ อีกมาก(กำหนด DEBUT 1 เมษายน 2569)