

การวิเคราะห์ถดถอยโลจิสติกแบบหลายกลุ่ม

Multinomial logistic regression analysis

คุณสมบัติของเทคนิควิธี

- ▶ ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหลายตัวที่เป็นข้อมูลเชิงปริมาณกับตัวแปรตาม 1 ตัว ที่เป็นข้อมูลเชิงคุณภาพที่แบ่งเป็นกลุ่ม (Categorical data) มากกว่า 2 กลุ่ม (Multicategorical)
 - ▶ นอกจากจะศึกษาผลโดยตรงของตัวแปรอิสระที่มีต่อตัวแปรตามแล้ว หากผู้วิจัยสนใจที่จะทดสอบ **สมมติฐานเกี่ยวกับผลของปฏิกริยา** (interaction effects) ที่มีต่อตัวแปรตาม และ **ความเป็นเส้นโค้ง** (Curvilinear effects) **ของความสัมพันธ์** ผู้วิจัยก็สามารถทำได้โดยใช้ผลคูณ (product terms) ของ 2 ตัวแปร หรือ 3 ตัวแปร ในการทดสอบผลของปฏิกริยาระหว่างตัวแปรอิสระ ที่มีต่อตัวแปรตาม และใช้ค่ายกกำลัง (Power term) เช่น ยกกำลังสอง หรือกำลังสาม ของตัวแปรอิสระแต่ละตัว เพื่อทดสอบความเป็นเส้นโค้ง โดยนำเอาผลคูณหรือผลยกกำลังเข้าไปในสมการวิเคราะห์แบบเชิงชั้น (Rhode, Orme, Cox and Bueher, 2003)
-



-
- ▶ ผลการวิเคราะห์จะชี้ให้เห็นว่าตัวแปรอิสระหรือตัวแปรพยากรณ์ (predictor) มีผลต่อตัวแปรตามหรือไม่ ทั้งในภาพรวม และของแต่ละตัวแปร ซึ่งพิจารณาได้จากค่าของ β และค่าอัตราส่วนโอกาส (Odd ratio) ผลที่ได้จากการวิเคราะห์ยังเปิดโอกาสให้ผู้วิจัยศึกษาเปรียบเทียบระหว่างกลุ่มย่อยๆ ว่าตัวแปรอิสระมีผลแตกต่างกันต่อทางเลือก (ความเป็นสมาชิกของกลุ่มย่อย) มากน้อยต่างกันอย่างไร โดยดูจากค่าอัตราส่วนโอกาส (Odd ratio) (Metsch, Pereya, Miles and McCoy, 2003)
 - ▶ เมื่อได้ผลการวิเคราะห์แล้ว ผู้วิจัยยังสามารถนำสมการที่ได้มานั้นมาประมาณค่าของโอกาสหรือค่าความน่าจะเป็นที่หน่วยวิเคราะห์แต่ละหน่วยจะตกอยู่ในกลุ่มย่อยกลุ่มใด (ของตัวแปรตาม) และดูค่าความถูกต้องของการจัดจำแนกแต่ละกลุ่ม และของทั้งหมดว่าเป็นเช่นใด ได้อีกด้วย
-

คุณสมบัติที่สำคัญของการวิเคราะห์ถดถอยมัลติโนเมียล

- ▶ เมื่อตัวแปรตามมี 3 กลุ่มย่อย สมการที่ใช้จะมี 2 สมการ
- ▶ เมื่อตัวแปรตามมี 4 กลุ่มย่อย สมการที่ใช้จะมี 3 สมการ

ทั้งนี้เพราะอีก 1 กลุ่ม ใช้เป็นกลุ่มอ้างอิง (reference group)

ตัวอย่าง การเลือกยานพาหนะในการเดินทาง สมมติว่ามีเพียง 3 ทางเลือก คือ รถประจำทาง (Bus) รถยนต์ส่วนตัว (Car) และ แท็กซี่ (Taxi) หากเราใช้ Taxi เป็นกลุ่มอ้างอิง เราจะได้ 2 สมการ ดังนี้

$$\log \left(\frac{P_{Bus}}{P_{Taxi}} \right) = \beta_0 + \beta_j X_j \quad (1)$$

$$\log \left(\frac{P_{Car}}{P_{Taxi}} \right) = \beta_0 + \beta_i X_i \quad (2)$$



-
- ▶ ในที่นี้ค่า P คือโอกาสหรือความน่าจะเป็นของทางเลือกนั้น β_0 คือจุดตัด (Intercept) β_i และ β_j ค่าสัมประสิทธิ์ถดถอย และ X_j และ X_i คือ ตัวแปรอิสระ ซึ่ง $i = j = n$ ตัวแปร
 - ▶ สมมติต่อจากตัวอย่างเดิมว่า มีทางเลือกเพิ่มอีก 1 ทาง เป็น 4 ทางเลือก คือเพิ่มไปทางเรือ (Boat) ผู้วิจัยจะต้องใช้ 3 สมการ โดย หากกลุ่มอ้างอิงคงเป็นการใช้รถแท็กซี่ สมการที่เพิ่มขึ้น คือ

$$\log \left(\frac{P_{Boat}}{P_{Taxi}} \right) = \beta_0 + \beta_k X_k \quad (3)$$



-
- ▶ สมการดังกล่าวอาจจะเขียนเป็นสมการทั่วไป ที่ระบุกลุ่มว่ามี j กลุ่ม มี i คน และตัวแปรอิสระมี 10 ตัว ได้ดังนี้

$$\ln \left(\frac{P(\text{category } i)}{P(\text{category } j)} \right) = \beta_{i0} + \beta_{i1} X_1 + \dots + \beta_{i10} X_{10} + \dots \quad (4)$$


ในสมการ (4) กลุ่มของตัวแปรตามมี j กลุ่ม โดยที่กลุ่มที่ j เป็นกลุ่มที่อ้างอิง ในที่นี้ $j=3$ (bus car และ taxi) โดย m เป็นกลุ่ม bus หรือ car และ j เป็นกลุ่ม taxi โดย X_1 ถึง X_{10} เป็นตัวแปรอิสระ ซึ่งบางตัวอาจจะเป็นตัวแปรหุ่นมีค่าเป็น 1 หรือ 0 ได้



-
- ▶ สมการทั้ง 3 สมการอาจเขียนสรุปเป็นสมการทั่วไป (general) สมการเดียวได้ ดังนี้

$$P(Y_i = m|X_i) = \frac{\exp(X_i\beta_m)}{1 + \sum_{j=2}^J \exp(X_i\beta_j)}$$

ในที่นี้ Y เป็นตัวแปรตามที่มีกลุ่มย่อย (สมมติว่า $j=3$) เมื่อ X_i เป็นเวกเตอร์ของตัวแปรอิสระ (ตัวพยากรณ์) ของคนที่ i และ β_m เป็นเวกเตอร์ของค่าสัมประสิทธิ์ของกลุ่มย่อยที่ m ความน่าจะเป็นหรือโอกาสของคนที่จะตกอยู่ในกลุ่มย่อยใด ประมาณค่าได้จาก $P(Y_i = m|X_i)$ ซึ่ง P คือทางเลือกที่เป็นไปได้ (Possible outcomes) และ $j = 1$ เป็นกลุ่มอ้างอิงในที่นี้คือการใช้รถแท็กซี่ และ $m > 1$

-
- ▶ ด้วยเหตุที่เทคนิคการวิเคราะห์นี้เป็นเทคนิคการวิเคราะห์ถดถอยพหุอย่างหนึ่ง ผู้วิจัยจึงอาจจะนำวิธีการหรือรูปแบบต่างๆ ของการวิเคราะห์ถดถอยพหุมาใช้ เช่น จะใช้วิธีการปกติ (ตัวแปรอิสระทุกตัวเข้าไปพร้อมกัน) หรือจะใช้วิธีการขั้นตอน (Stepwise) หรือ จะใช้วิธีการแบบเชิงชั้น (Hierarchical)
 - ▶ ความแตกต่างในด้านประโยชน์และผลที่ได้จากการวิเคราะห์เป็นไปตามวัตถุประสงค์หรือเหตุผลความจำเป็นในการเลือกวิธีนั้นๆ
-
- 

ความต้องการด้านข้อมูล

- ▶ ตัวแปรตามต้องเป็นข้อมูลเชิงคุณภาพ คือ เป็นตัวแปรกลุ่มที่มีหลายกลุ่มมากกว่า 2 กลุ่มขึ้นไป (Multicategorical)
- ▶ ตัวแปรอิสระต้องเป็นข้อมูลเชิงปริมาณ หรือตัวแปรเชิงคุณภาพ



ตัวอย่างงานวิจัยที่ 1

- ▶ Akin, J.S., Schwartz, J.B. (1988). The Effect of economic factor on contraceptive choice in Jamaica and Thailand: A comparison of mixed multinomial logit results. *Economic Development and Cultural Change*, 36: 134-153
 - ▶ เกี่ยวกับการตัดสินใจเลือกวิธีการคุมกำเนิดซึ่งมี 6 วิธีคือ ถุงยางอนามัย ห่วง ยาฉีด ยาคุม ทำหมัน วิธีธรรมชาติ
 - ▶ ตัวแปรอิสระแบ่งออกเป็น 2 กลุ่ม
 - ▶ ตัวแปรเงื่อนไข (Conditional variables) ได้แก่ ราคาของการคุมครอง 1 เดือนของแต่ละวิธี เวลาที่ต้องใช้เดินทาง อัตราประสิทธิผลของแต่ละวิธี
 - ▶ ตัวแปรที่ไม่เป็นเงื่อนไข (Unconditional variables) คุณสมบัติเฉพาะบุคคล เช่น รายได้ การศึกษา อายุ เพศ เป็นต้น
-



คุณสมบัติของตัวแปรอิสระที่ใช้ในการวิเคราะห์ Logistic Regression

Description of Variables, Thailand Binary Logit Equation

CONDP	= Price of one month's protection using condoms, in dollars
IUDP	= Price of one month's protection using IUD, in dollars
INJP	= Price of one month's protection using injection, in dollars
PILLP	= Price of one month's protection using pills, in dollars
STERP	= Price of one month's protection using sterilization, in dollars
CONDTP	= Time required per month to travel and wait to purchase condoms, weighted by household income
IUDTP	= Time required per month to travel and wait to purchase IUD, weighted by household income
INJTP	= Time required per month to travel and wait to purchase injection, weighted by household income
PILLTP	= Time required per month to travel and wait to purchase pills, weighted by household income
STERTP	= Time required per month to travel and wait to purchase sterilization, weighted by household income
CONDTPF	= Time required per month to travel and wait for free condoms, weighted by household income
IUDTPF	= Time required per month to travel and wait for free IUD, weighted by household income



INJTPF	= Time required per month to travel and wait for free injections, weighted by household income
PILLTPF	= Time required per month to travel and wait for free pills, weighted by household income
STERTPF	= Time required per month to travel and wait for free sterilization, weighted by household income
INCOME	= Monthly household income, in hundreds of dollars
EDUC	= Respondent's education level: 1 = no formal education; 2 = primary school; . . . ; 8 = bachelor's degree or higher
URBAN	= 1 = urban residence; 0 = rural
AGE	= Respondent's age in years
BUDHA	= 1 = respondent's religion is Buddhist; 0 = other
EXEC	= 1 = executive occupation; 0 = other
BUS	= 1 = business occupation; 0 = other
SALES	= 1 = sales occupation; 0 = other
SKILL	= 1 = skilled occupation; 0 = other
BLUE	= 1 = blue-collar occupation; 0 = other



DESCRIPTIVE STATISTICS, THAILAND BINARY LOGIT EQUATION

Independent Variable	Mean	Standard Deviation	Minimum	Maximum
CONDP	2.220	.902	.286	14.286
IUDP	.304	.132	.015	2.083
INJP	.589	.207	.060	2.976
PILLP	.529	.222	.036	2.964
STERP	.162	.063	.006	.893
CONDTP	43.856	48.363	1.339	580.357
IUDTP	7.053	6.898	.112	78.125
INJTP	36.016	43.327	.149	580.357
PILLTP	23.566	32.392	.297	580.350
STERTP	.828	1.617	.007	35.714
CONDTPF	62.768	64.108	1.786	745.134
IUDTPF	10.659	10.838	.149	127.143
INJTPF	56.892	57.526	1.339	675.848
PILLTPF	41.359	43.991	.298	496.629
STERTPF	.938	1.023	.022	10.915
INCOME	1.872	1.866	.179	22.321
EDUC	2.953	1.724	1.000	8.000
URBAN	.722	.448	0	1.000
AGE	31.759	6.739	15.000	44.000
BUDHA	.969	.389	0	1.000
EXEC	.081	.273	0	1.000
BUS	.191	.393	0	1.000
SALES	.205	.404	0	1.000
SKILL	.140	.348	0	1.000
BLUE	.178	.308	0	1.000

BINARY LOGIT RESULTS FOR THAILAND—DEPENDENT VARIABLE: PROBABILITY OF USING A
CONTRACEPTIVE METHOD NOT SUPPLIED FREE OF CHARGE

Independent Variable	Coefficient	Asymptotic <i>t</i> -Value
Constant	-.905	-1.107
CONDP	.003	.034
IUDP	-.207	-.318
INJP	.016	.040
PILLP	-.541	-1.322
STERP	-1.195	-.883
CONDTP	-.001	-.203
IUDTP	-.010	-.407
INJTP	.002	.318
PILLTP	.003	.503
STERTP	.068	.549
CONDTPF	-.002	-.462
IUDTPF	.041	.891
INJTPF	-.005	-.485
PILLTPF	.004	.725
STERTPF	-.097	-1.076
INCOME	.211	.660
EDUC	.041	.623
URBAN	.384***	1.770
AGE	.028**	2.103
BUDHA	.361	.731
EXEC	-.166	-.379
BUS	.807**	2.067
SALES	.483	1.393
SKILL	.539	1.527
BLUE	.362	1.144

NOTES.—No. of observations, dependent variable = 0: 228; no. of observations, dependent variable = 1: 497; total observations: 725; $\chi^2_{25} = 47.404$.

** Significant at $\alpha = .05$.

*** Significant at $\alpha = .10$.

Multinomial Logit Model

Description of Variables, Thailand Multinomial Logit Model

- PRICE** = Price of one month's protection for each of the six possible methods, in dollars: condom, IUD, injection, pill, natural and sterilization
- TIME** = Time required per month to travel and wait for each of the six possible methods, weighted by household income
- EFFECT** = Perfect effectiveness for each of the six possible methods
- INCOME** = Monthly household income, in hundreds of dollars
- EDUC** = Respondent's education level
- AGE** = Respondent's age, in years
- URBAN** = 1 = urban residence; 0 = rural
- EXEC** = 1 = executive; 0 = other
- BUS** = 1 = business occupation; 0 = other
- SALES** = 1 = sales occupation; 0 = other
- SKILL** = 1 = skilled occupation; 0 = other
- MCHILD** = 1 = does not desire more children; 0 = desires more children
-



ค่าเฉลี่ยของตัวแปรอิสระ จำแนกตามวิธีการคุมกำเนิดที่ใช้ในสมการวิเคราะห์
Multinomial Logit

INDEPENDENT VARIABLE MEANS, THAILAND MULTINOMIAL LOGIT EQUATION

Independent Variable	Condom	IUD	Injection	Pill	Natural Method	Sterilization	All
PRICE	2.224	.305	.594	.530	0	.160	.635
TIME	57.343	7.802	39.972	26.952	0	.950	22.182
EFFECT	.900	.950	.998	.980	.760	.999	.231
INCOME	3.530	2.429	1.465	1.652	2.422	1.956	2.018
EDUC	4.923	3.600	2.565	3.134	3.438	2.950	3.099
AGE	28.346	30.800	30.275	29.465	30.938	35.201	31.903
URBAN	.962	.867	.652	.767	.813	.804	.781
EXEC	.231	.133	.145	.523	.125	.125	.091
BUS	.231	.200	.870	.186	.188	.231	.193
SALES	.192	.133	.232	.279	.250	.236	.245
SKILL	.190	.133	.232	.093	.250	.166	.153
MCHILD	.500	.600	.637	.955	.563	1.000	.748

สถิติพรรณนาตัวแปรอิสระที่ใช้ในสมการ Multinomial Logit

DESCRIPTIVE STATISTICS, THAILAND MULTINOMIAL LOGIT EQUATION

Independent Variable	Mean	Standard Deviation	Minimum	Maximum
CONDP*	2.224	1.085	.286	14.286
IUDP*	.305	.159	.015	2.083
INJP*	.594	.249	.059	2.976
PILLP*	.530	.267	.036	2.964
STERP*	.160	.075	.006	.893
CONDTP*	57.343	62.848	1.339	658.973
IUDTP*	7.802	7.937	.112	78.125
INJTP*	39.972	50.656	.149	580.357
PILLTP*	26.952	38.698	.298	580.350
STERTP*	.950	1.937	.007	35.714
INCOME	2.018	2.102	.179	22.321
EDUC	3.099	1.797	1.000	8.000
AGE	31.903	6.841	15.000	44.000
URBAN	.781	.419	0	1.000
EXEC	.091	.271	0	1.000
BUS	.193	.408	0	1.000
SALES	.245	.412	0	1.000
SKILL	.153	.361	0	1.000
MCHILD	.748	.422	0	1.000

ผลการวิเคราะห์ Multinomial Logit

MULTINOMIAL LOGIT RESULTS, THAILAND

Independent Variable	Coefficient	Asymptotic t-Value
Conditional variables:		
PRICE	-.322	-1.559
TIME	-.030*	-5.748
EFFECT	3.623	.618
Unconditional variables:		
Log $\left[\frac{PCOND}{PPILL} \right]$:		
INCOME	.129	1.084
EDUC	.272**	1.903
AGE	-.100*	-2.776
URBAN	1.322	1.410
EXEC	1.051	.933
BUS	.379	.448
SALES	-.561	-.668
SKILL	1.237***	1.669
MCHILD	-.347	-.639
Log $\left[\frac{PIUD}{PPILL} \right]$:		
INCOME	-.154	-1.145
EDUC	-.054	-.312
AGE	-.070*	-2.587
URBAN	.346	.484
EXEC	.779	.690

Independent Variable	Coefficient	Asymptotic t-Value
BUS	-.421	-.482
SALES	-1.239	-1.345
SKILL	-.154	-.172
MCHILD	-.236	-.383
Log $\left[\frac{\text{PINJ}}{\text{PPILL}} \right]$:		
INCOME	-.010	-.097
EDUC	-.155	-1.342
AGE	.008	.505
URBAN	-.623***	-1.681
EXEC	-.283	-.241
BUS	-.133	-.225
SALES	.215	.466
SKILL	1.121**	2.399
MCHILD	-.270	-.699
Log $\left[\frac{\text{PSTER}}{\text{PPILL}} \right]$:		
INCOME	-.387*	-3.715
EDUC	-.302*	-3.736
AGE	.030*	2.693
URBAN	-.422	-1.352
EXEC	2.247*	3.888
BUS	1.138*	2.977
SALES	.708**	2.062
SKILL	.935**	2.394
MCHILD	1.336*	4.091
Log $\left[\frac{\text{PNAT}}{\text{PPILL}} \right]$:		
INCOME	-.214	-1.337
EDUC	-.195	-1.100
AGE	-.042	-.973
URBAN	-.625	-.836
EXEC	1.986***	1.756
BUS	.768	.769
SALES	.701	.790
SKILL	1.613***	1.869
MCHILD	-.515	-.834

NOTE:—No. of observations: for condoms: 26; for IUDs: 15; for injections: 69; for the pill: 172; for natural method: 16; for sterilization: 199; total: 497. $\chi^2_{43} = 608.080$. PCOND = probability of using condom; PIUD = probability of using IUD; PINJ = probability of using injection; PPILL = probability of using the pill; PNAT = probability of using natural method; PSTER = probability of using sterilization.

* Significant at $\alpha = .01$.

** Significant at $\alpha = .05$.


*** Significant at $\alpha = .10$.

ผลของการวิเคราะห์

- ▶ เทคนิคการวิเคราะห์โลจิสติกส์โนเมียลแบบผสมมีเงื่อนไข และไม่มีเงื่อนไข
- ▶ ผลที่ได้ชี้ให้เห็นว่าในกลุ่มสตรีที่ไม่ได้เลือกใช้วิธีการจากแหล่งให้เปล่า ราคาเป็นปัจจัยสำคัญและมีความสัมพันธ์ในเชิงลบกับวิธีคุมกำเนิดที่เลือกใช้ ซึ่งสามารถตีความหมายได้ว่าเมื่อราคาการป้องกันต่อเดือนของวิธีใดก็ตามที่สูงขึ้น คนก็จะหันไปใช้วิธีอื่น สำหรับกลุ่มผู้ซื้อการคุมกำเนิดราคาที่สูงขึ้นจะมีแนวโน้มที่จะลดโอกาสที่จะเลือกใช้วิธีดังกล่าว
- ▶ ค่าของเวลาที่ใช้ในการเดินทางและรอรับบริการมีนัยสำคัญและติดเครื่องหมายลบ ซึ่งหมายความว่า ถ้าค่าของเวลาของวิธีการคุมกำเนิดสูงขึ้น โอกาสที่จะเลือกใช้วิธีการนั้นจะลดลง



สำหรับการตีความหมายผลของตัวแปรที่ไม่ใช่เงื่อนไข จะต้องเป็นไปในเชิงสัมพันธ์กับกลุ่มฐาน (base category) ซึ่งในที่นี้ได้แก่ ยาเม็ดคุมกำเนิด ผลที่เสนอของแต่ละวิธีตัวเลือกจึงเปรียบเทียบกับยาเม็ดคุมกำเนิด

- ▶ ตัวแปรตาม $\log \left(\frac{PCOND}{PPILL} \right)$ และตัวแปรอิสระ คือรายได้ (income) การศึกษา (educ) อายุ (Age) เขตเมือง (Urban) ผู้บริหาร (Exec) ธุรกิจ (Bus) พนักงานขาย (sales) แรงงานฝีมือ (Skill) และความต้องการบุตรเพิ่ม (mchild) ซึ่งพบว่ารายได้ อายุ และแรงงานฝีมือมีความสัมพันธ์อย่างมีนัยสำคัญต่อการเลือกใช้ถุงยางอนามัยเมื่อเปรียบเทียบกับการใช้ยาเม็ดคุมกำเนิด
-
- 

การพยากรณ์ค่าของโอกาสที่จะเลือกการคุมกำเนิดแต่ละวิธี

วิธี	ค่าของโอกาส(ความน่าจะเป็น)
ถุงยางอนามัย	0.0202
ยาฉีด	0.1049
ยาเม็ด	0.3518
ห่วง	0.0273
ทำหมัน	0.4609
วิธีธรรมชาติ	0.0349

หมายเหตุ ค่าของโอกาสคาดประมาณจากตัวแปรอิสระทุกตัว ณ ค่าเฉลี่ย ของแต่ละตัว



ตัวอย่างงานวิจัยที่ 2

- ▶ Mahapatra, K. & Kent, S. (2003). Tropical deforestation: A multinomial logistic model and some country-specific policy prescriptions. *Forest Policy and Economics*.
 - ▶ การสูญเสียป่าเมืองร้อน: การสร้างแบบจำลอง Multinomial และนโยบายการแก้ไขเฉพาะประเทศ
 - ▶ ข้อมูล 117 ประเทศ ใน Africa Asia Latin America จาก FRA
 - ▶ ผู้วิจัยทำการศึกษาเปรียบเทียบการใช้แบบจำลอง Multinomial, Binomial และ การวิเคราะห์ OLS
 - ▶ ผู้วิจัยได้ตระหนักถึงปัญหา Heteroscedasticity, Linearity
-



$$DEF = f(FORAR, POPGR, GDPGR, DEBTGR, AGGR, ROAD, DEMOCRACY) + e$$

$$\ln \left(\frac{P(\text{category}_i)}{P(\text{category}_j)} \right) = \beta_{i0} + \beta_{i1} X_1 + \beta_{i2} X_2 + \dots + \beta_{i10} X_{10} + e$$

- ▶ ในการวิเคราะห์ Multinomial ตัวแปรตาม คือ การสูญเสีย/ตัดไม้ทำลายป่า ออกเป็น 3 กลุ่ม คือ ระดับสูง ระดับกลาง ระดับต่ำ โดยให้ระดับต่ำเป็นกลุ่มอ้างอิง
 - ▶ ส่วนแบบจำลอง Binomial แบ่งอัตราการสูญเสียป่าไม้ออกเป็น 2 กลุ่ม สูง และ ต่ำ เป็นตัวแปรตาม
 - ▶ การวิเคราะห์ OLS ผู้วิจัยใช้อัตราการสูญเสียป่าไม้เป็นข้อมูลต่อเนื่อง
-



▶ ผลการวิเคราะห์ Multinomial

▶ อัตราส่วนความเป็นไปได้ (Likelihood ratio) ของตัวแปรแต่ละตัวมีนัยสำคัญทางสถิติอย่างน้อยระดับที่ 0.10

▶ สถิติทดสอบ Wald เกี่ยวกับนัยสำคัญทางสถิติของค่าสัมประสิทธิ์ชี้ให้เห็นว่าตัวแปรอิสระ

▶ *FORAR, POPGR, AGGR, ROAD, IDASIA* และ *IDLAT* มีนัยสำคัญทางสถิติกับค่า Logit ของ High/Low

▶ *DEBTGR, AGGR, ROAD* มีนัยสำคัญทางสถิติกับค่า Logit ของ Med/Low



- ▶ ผลการวิเคราะห์ Binomial
- ▶ เครื่องหมาย +/- ของทุกตัวแปรในแบบจำลอง High/Low ของ Multinomial Logistic เหมือนกับของ แบบจำลอง High/Low ของ Binary Logistic แต่ เครื่องหมายของตัวแปร ***GDPGR*** และ ***IDASIA*** ใน Med/Low ของโลจิท ในแบบจำลอง Multinomial ตรงกันข้ามกับเครื่องหมายของตัวแปรเหล่านี้ใน แบบจำลอง Binary Logistic
- ▶ ตัวแปร ***AGGR, ROAD, IDASIA*** มีนัยสำคัญ และตัวแปร ***GDPGR*** และ ***DEMOCRACY*** ไม่มีนัยสำคัญในแบบจำลอง Binary Logistic และ แบบจำลอง Multinomial
- ▶ ***POPGR*** มีนัยสำคัญในแบบจำลอง Binary Logistic ซึ่งบ่งชี้ถึงความเป็น ปัจจัยเชิงสาเหตุ (casual factor) ของการตัดไม้ทำลายป่าในประเทศที่มีการ สูญเสียป่าไม้ระดับสูง



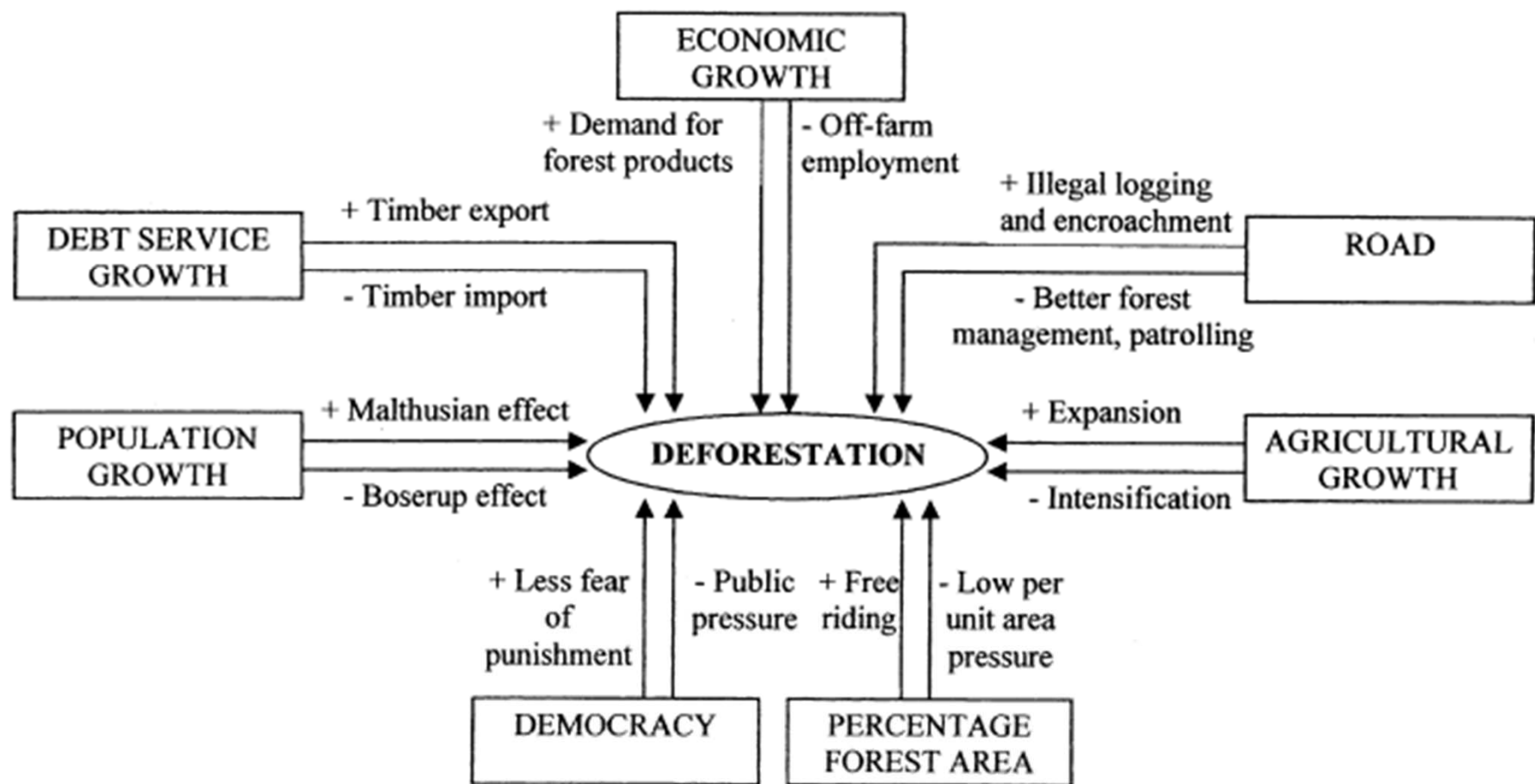


Fig. 1. Dual effects of the explanatory variables on deforestation.

รายละเอียดของข้อมูลและการให้ค่านิยามนามตัวแปรที่ใช้ในการวิเคราะห์

Details of the data used for the estimation of deforestation model

Sector	Variables	Explanation	Unit	Source
Forest	DEF	Average annual rate of deforestation from 1980–1990 and 1990–1995	Percent	FAO, 1993 FAO, 1997
	FORAR	Percentage of land area of a country under forests in 1980 and 1990	Percent	FAO, 1993 FAO, 1997
Demographic	POPGR	Average annual growth rate of population from 1980–1990 and 1990–1995	Percent	FAO, 1993 FAO, 1997
Macro-economic	GDPGR*	Average annual growth rate of GDP (excluding agriculture) during 1980–1990 and 1990–1995	Percent	World Bank, 1999
	DEBTGR*	Average annual growth rate of Total Debt service as a percentage of GNP during 1980–1990 and 1990–1995	Percent	World Bank, 1999
Agriculture	AGGR	Average annual growth rate of agricultural sector during 1980–1990 and 1990–1995	Percent	World Bank, 1997
Infrastructure	ROAD	Percentage paved road of a country for the years 1990 and 1995	Percent	World Bank, 1999
Political	DEMOCRACY	Average DEMOCRACY index of a country during 1980–1990 and 1990–1995	Integer	Gurr and Jagers, 1999

Note: *Values of these variables are calculated.



สถิติพรรณนาตัวแปรอิสระที่ใช้ในแบบจำลองการสูญเสียป่าไม้

Descriptive statistics for the variables of the deforestation model ($N=117$)

	DEF	FORAR	POPGR	GDPGR	DEBTGR	AGGR	ROAD	DEMOCRACY
Mean	1.11	33.85	2.65	2.77	1.82	2.34	21.38	3.47
Min.	0.00	0.47	0.90	-12.95	-55.11	-10.80	0.17	0.00
Max.	3.90	95.48	4.20	10.29	37.45	12.20	75.00	10.00
S.D.	0.91	21.84	0.65	3.65	12.47	2.89	15.77	3.55



ผลการวิเคราะห์ Multinomial Logit

Results of multinomial logistic regression ($N=117$)

Variable	Likelihood Ratio test Chi-square	Wald test					
		High/Low			Med/Low		
		β	Wald	Exp (β)	β	Wald	Exp ()
INTERCEPT	13.592*	-6.670	10.695*		-4.084	4.893*	
FORAR	4.315 ⁺	-0.0299	3.696*	0.971	-0.00375	0.082	0.996
POPGR	7.116*	1.420	6.194*	4.138	0.791	2.345	2.206
GDPGR	2.660	-0.178	2.016	0.837	0.02285	0.055	1.023
DEBTGR	7.564*	0.01199	0.212	1.012	0.06147	6.508*	1.063
AGGR	7.801*	0.263	2.855*	1.301	0.311	5.947*	1.365
ROAD	7.991*	0.04206	4.348*	1.043	0.05161	5.848*	1.053
DEMOCRACY	1.085	0.03812	0.146	1.039	0.09467	1.055	1.099
IDASIA	25.352*	3.745	10.298*	42.326	-2.039	2.907*	0.130
IDLAT	29.171*	4.511	16.522*	91.018	0.08741	0.008	1.091

Note: *Significant at 10% level of significance. The critical values at 10% level of significance are 2.705 for 1 degree of freedom (Wald test) and 4.605 for 2 degrees of freedom (LR test).

⁺Significant at 11.6% level of significance (significance of this result is discussed in Section 8).



ตัวอย่างงานวิจัยที่ 3

- ▶ Peng, C.J. & Nichols, R.N. (2003). Using multinomial logistic models to predict adolescent behavioral risk. *Journal of Modern Applied Statistical Methods*. 2(1): 1-13.
- ▶ การใช้แบบจำลอง Multinomial Logistic ในการพยากรณ์พฤติกรรมเสี่ยงของวัยรุ่น
- ▶ ตัวอย่าง เด็กวัยรุ่นเรียนมัธยมศึกษาจำนวน 432 คน คณะผู้วิจัยได้สร้างมาตรฐานวัดพฤติกรรมเสี่ยงด้านสุขภาพ แบ่งพฤติกรรมเสี่ยงออกเป็น 3 กลุ่ม โดยใช้ค่าเบี่ยงเบนมาตรฐานจากค่าเฉลี่ยเป็นจุดแบ่งกลุ่มที่มีพฤติกรรมเสี่ยงสูงสุด ปานกลาง และต่ำ
- ▶ ตัวแปรอิสระ- เพศ ประเภทครอบครัว ความเสี่ยงทางอารมณ์



การวิเคราะห์ Multinomial Logit พฤติกรรมเสี่ยงต่อการทำร้ายตนเองของวัยรุ่น

Predictor	β	<i>SE</i> β	Wald's χ^2 (<i>df</i> =1)	<i>p</i>	e^β (odds ratio)
CONSTANT 1 (Y_1)	-0.6211	1.0627	0.3416	0.5589	Not necessary
CONSTANT 2 (Y_1+Y_2)	2.5220	1.0723	5.5317	0.0187	Not necessary
<u>GENDER</u> (boys=1, girls=0)	1.1070	0.2111	27.5060	<0.0001	3.0253
<u>DROPOUT</u> (yes=1, no=0)	2.1818	0.3287	44.0618	<0.0001	8.8622
<u>FAMILY</u>	0.4135	0.1179	12.2979	<0.001	1.5121
<u>EMOTION</u>	0.0074	0.0115	0.4118	0.5211	1.0074
<u>ESTEEM</u>	-0.0488	0.0118	16.9867	<0.0001	0.9524
<u>Overall Model Evaluation</u>					
Tests	χ^2	<i>df</i>	<i>p</i>		
Likelihood Ratio Test	122.38	5	<0.0001		
Score test	110.47	5	<0.0001		
Wald test	97.87	5	<0.0001		

Notes. Cox and Snell *R* squared=0.2467. Nagelkerke *R* squared (Max rescaled *R* squared)=0.2978. Kendall's Tau-*a* = 0.297. Goodman-Kruskal's Gamma= 0.548. Somers' D_{xy} = 0.539. *c*-statistic = 0.769.

Table 3. Predicated Probability of Participating in Self-injurious Behavior for 36 Children.

Case No.	SEX $\beta=1.107$ 1=boy 0=girl	DROPOUT $\beta=2.1818$ 1=yes 0=no	FAMILY $\beta=0.4135$ 1=intact, 2=step, 3=single	EMOTION $\beta=0.0074$	ESTEEM $\beta=-0.0488$	Intercept 1 $\alpha_1=-0.6211$	Intercept 2 $\alpha_2=2.522$	Predicted probability of participating in self-injurious behavior			Actual Behavior risk, 1=high, 2=med, 3=low (score on HBO, M=47.69, SD=10.89)
								p_1 (high)	p_2 (medium)	p_3 (low)	
1	1	0	1	62.39	32.68	-0.6211	2.5220	.0818	.5921	.3261	1 (60.40)
2	1	0	1	80.74	32.68	-0.6211	2.5220	.0926	.6102	.2972	2 (52.77)
3	1	0	1	32.07	71.58	-0.6211	2.5220	.0106	.1878	.8016	3 (42.65)
4	1	1	1	72.72	46.41	-0.6211	2.5220	.3038	.6062	.0900	1 (95.21)
5	1	1	1	63.07	37.25	-0.6211	2.5220	.3885	.5479	.0636	2 (50.00)
6	1	1	1	---	---	-0.6211	2.5220	---	---	---	3 (-----)
7	0	0	1	47.29	41.83	-0.6211	2.5220	.0166	.2645	.7189	1 (61.53)
8	0	0	1	45.78	44.12	-0.6211	2.5220	.0147	.2422	.7431	2 (47.07)
9	0	0	1	42.05	21.24	-0.6211	2.5220	.0425	.4643	.4932	3 (42.70)
10	0	1	1	51.37	34.97	-0.6211	2.5220	.1772	.6559	.1669	1 (70.23)
11	0	1	1	56.77	37.25	-0.6211	2.5220	.1670	.6559	.1771	2 (53.27)
12	0	1	1	---	---	-0.6211	2.5220	---	---	---	3 (-----)
13	1	0	2	41.36	50.98	-0.6211	2.5220	.0451	.4776	.4773	1 (72.83)
14	1	0	2	46.14	50.98	-0.6211	2.5220	.0467	.4848	.4685	2 (45.84)
15	1	0	2	36.11	41.83	-0.6211	2.5220	.0663	.5559	.3778	3 (40.44)
16	1	1	2	38.59	57.85	-0.6211	2.5220	.2269	.6449	.1282	1 (92.50)
17	1	1	2	54.87	46.41	-0.6211	2.5220	.3665	.5641	.0694	2 (46.99)
18	1	1	2	70.35	34.97	-0.6211	2.5220	.5312	.4321	.0367	3 (43.52)
19	0	0	2	---	---	-0.6211	2.5220	---	---	---	1 (-----)
20	0	0	2	34.21	44.12	-0.6211	2.5220	.0203	.3041	.6756	2 (45.78)
21	0	0	2	50.18	53.27	-0.6211	2.5220	.0147	.2421	.7432	3 (40.44)
22	0	1	2	---	---	-0.6211	2.5220	---	---	---	1 (-----)
23	0	1	2	54.84	50.98	-0.6211	2.5220	.1326	.6473	.2201	2 (48.64)
24	0	1	2	50.18	46.41	-0.6211	2.5220	.1559	.6547	.1894	3 (43.08)
25	1	0	3	63.52	23.52	-0.6211	2.5220	.2432	.6384	.1184	1 (67.90)
26	1	0	3	32.07	67.00	-0.6211	2.5220	.0296	.3848	.5856	2 (56.69)
27	1	0	3	50.18	48.70	-0.6211	2.5220	.0786	.5854	.3360	3 (40.44)
28	1	1	3	43.54	48.70	-0.6211	2.5220	.4184	.5250	.0566	1 (85.49)
29	1	1	3	56.74	44.12	-0.6211	2.5220	.4979	.4604	.0417	2 (54.31)
30	1	1	3	---	---	-0.6211	2.5220	---	---	---	3 (-----)
31	0	0	3	---	---	-0.6211	2.5220	---	---	---	1 (-----)
32	0	0	3	64.12	28.10	-0.6211	2.5220	.0786	.5856	.3358	2 (48.41)
33	0	0	3	60.08	39.54	-0.6211	2.5220	.0453	.4781	.4766	3 (44.41)
34	0	1	3	---	---	-0.6211	2.5220	---	---	---	1 (-----)
35	0	1	3	43.63	48.70	-0.6211	2.5220	.1922	.6543	.1535	2 (46.34)
36	0	1	3	---	---	-0.6211	2.5220	---	---	---	3 (-----)

การอ่านผล

- ▶ ภาพรวมของแบบจำลอง

- ▶ ค่าอัตราส่วนความเป็นไปได้ (likelihood ratio)
- ▶ คะแนน (Score)
- ▶ ค่าสถิติ Wald

ยืนยันว่าแบบจำลองมีประสิทธิภาพมากกว่าแบบจำลองว่าง หรือแบบจำลองที่มีจุดตัดแต่เพียงอย่างเดียวโดยไม่มีตัวแปรอิสระ

- ▶ สถิติทดสอบตัวแปรอธิบายแต่ละตัว

- ▶ ค่าสัมประสิทธิ์ถดถอยที่ทดสอบด้วยค่าสถิติ Wald พบว่าตัวแปรทุกตัว ยกเว้นตัวแปรความเสี่ยงทางอารมณ์มีผลต่อพฤติกรรมเสี่ยงของวัยรุ่น และพบว่าค่าคงที่ ทั้งสองตัวมีนัยสำคัญ จึงควรรวมอยู่ในแบบจำลองด้วย



-
- ▶ สถิติทดสอบความเหมาะสมของแบบจำลอง
 - ▶ ค่า R ยกกำลังของ Cox and Snell
 - ▶ การตรวจสอบความถูกต้องของการประมาณค่าโอกาส/ความน่าจะเป็น
 - ▶ คณะผู้วิจัยได้ใช้แบบจำลองที่ได้จากการวิเคราะห์มาคาดประมาณค่าโลจิทของพฤติกรรมเสี่ยงระดับกลางและระดับสูง



ตัวอย่าง STATA

- ▶ ตัวแปรตาม ประเภทของประกันสุขภาพ

insure	Freq.	Percent	Cum.
Indemnity	294	47.73	47.73
Prepaid	277	44.97	92.69
Uninsure	45	7.31	100.00
Total	616	100.00	

insure	nonwhite		Total
	0	1	
Indemnity	251 50.71	43 35.54	294 47.73
Prepaid	208 42.02	69 57.02	277 44.97
Uninsure	36 7.27	9 7.44	45 7.31
Total	495 100.00	121 100.00	616 100.00

Pearson chi2(2) = 9.5599 Pr = 0.00



Multinomial logistic regression

Number of obs = 616

LR chi2(2) = 9.62

Prob > chi2 = 0.0081

Log likelihood = -551.78348

Pseudo R2 = 0.0086

insure	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Indemnity	(base outcome)					
Prepaid						
nonwhite	.6608212	.2157321	3.06	0.002	.2379942	1.083648
_cons	-.1879149	.0937644	-2.00	0.045	-.3716896	-.0041401
Uninsure						
nonwhite	.3779586	.407589	0.93	0.354	-.4209011	1.176818
_cons	-1.941934	.1782185	-10.90	0.000	-2.291236	-1.592632



Multinomial logistic regression

Number of obs = 616

LR chi2(2) = 9.62

Prob > chi2 = 0.0081

Log likelihood = -551.78348

Pseudo R2 = 0.0086

insure	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Indemnity						
nonwhite	-.6608212	.2157321	-3.06	0.002	-1.083648	-.2379942
_cons	.1879149	.0937644	2.00	0.045	.0041401	.3716896
Prepaid (base outcome)						
Uninsure						
nonwhite	-.2828627	.3977302	-0.71	0.477	-1.0624	.4966742
_cons	-1.754019	.1805145	-9.72	0.000	-2.107821	-1.400217



Odd ratio

Multinomial logistic regression

Number of obs = 616

LR chi2(2) = 9.62

Prob > chi2 = 0.0081

Log likelihood = -551.78348

Pseudo R2 = 0.0086

insure	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
Indemnity						
nonwhite	.516427	.1114099	-3.06	0.002	.3383588	.7882073
_cons	1.206731	.1131483	2.00	0.045	1.004149	1.450183
Prepaid (base outcome)						
Uninsure						
nonwhite	.7536233	.2997387	-0.71	0.477	.3456255	1.643247
_cons	.1730769	.0312429	-9.72	0.000	.1215024	.2465434



Multinomial logistic regression

Number of obs = 615
 LR chi2(10) = 42.99
 Prob > chi2 = 0.0000
 Pseudo R2 = 0.0387

Log likelihood = -534.36165

insure	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Indemnity (base outcome)						
Prepaid						
age	-.011745	.0061946	-1.90	0.058	-.0238862	.0003962
male	.5616934	.2027465	2.77	0.006	.1643175	.9590693
nonwhite	.9747768	.2363213	4.12	0.000	.5115955	1.437958
site						
2	.1130359	.2101903	0.54	0.591	-.2989296	.5250013
3	-.5879879	.2279351	-2.58	0.010	-1.034733	-.1412433
_cons	.2697127	.3284422	0.82	0.412	-.3740222	.9134476
Uninsure						
age	-.0077961	.0114418	-0.68	0.496	-.0302217	.0146294
male	.4518496	.3674867	1.23	0.219	-.268411	1.17211
nonwhite	.2170589	.4256361	0.51	0.610	-.6171725	1.05129
site						
2	-1.211563	.4705127	-2.57	0.010	-2.133751	-.2893747
3	-.2078123	.3662926	-0.57	0.570	-.9257327	.510108
_cons	-1.286943	.5923219	-2.17	0.030	-2.447872	-.1260134

Multinomial logistic regression

Number of obs = 615

LR chi2(10) = 42.99

Prob > chi2 = 0.0000

Log likelihood = -534.36165

Pseudo R2 = 0.0387

insure	RRR	Std. Err.	z	P> z	[95% Conf. Interval]	
Indemnity (base outcome)						
Prepaid						
age	.9883237	.0061223	-1.90	0.058	.9763968	1.000396
male	1.75364	.3555444	2.77	0.006	1.178588	2.609267
nonwhite	2.650576	.6263875	4.12	0.000	1.66795	4.212086
site						
2	1.119672	.2353442	0.54	0.591	.7416116	1.690461
3	.5554437	.1266051	-2.58	0.010	.3553214	.868278
_cons	1.309588	.430124	0.82	0.412	.6879617	2.492902
Uninsure						
age	.9922342	.011353	-0.68	0.496	.9702304	1.014737
male	1.571216	.5774008	1.23	0.219	.7645935	3.228799
nonwhite	1.242417	.5288177	0.51	0.610	.5394676	2.861341
site						
2	.2977316	.1400865	-2.57	0.010	.1183924	.7487316
3	.8123595	.2975613	-0.57	0.570	.396241	1.665471
_cons	.2761136	.1635481	-2.17	0.030	.0864774	.881603

Test Prepaid

```
( 1) [Prepaid]age = 0
( 2) [Prepaid]male = 0
( 3) [Prepaid]nonwhite = 0
( 4) [Prepaid]1b.site = 0
( 5) [Prepaid]2.site = 0
( 6) [Prepaid]3.site = 0
      Constraint 4 dropped

           chi2( 5) =    28.84
Prob > chi2 =    0.0000
```



Test Uninsure

```
( 1) [Uninsure]age = 0
( 2) [Uninsure]male = 0
( 3) [Uninsure]nonwhite = 0
( 4) [Uninsure]1b.site = 0
( 5) [Uninsure]2.site = 0
( 6) [Uninsure]3.site = 0
      Constraint 4 dropped
```

```
      chi2( 5) =      9.31
Prob > chi2 =      0.0973
```



หนังสืออ้างอิง

- ▶ สุขชาติ ประสิทธิ์รัฐสินธุ์ (2551) เทคนิคการวิเคราะห์ตัวแปรหลายตัวสำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์. หจก. สามลดา
 - ▶ Akin, J.S., Schwartz, J.B. (1988). The Effect of economic factor on contraceptive choice in Jamaica and Thailand: A comparison of mixed multinomial logit results. *Economic Development and Cultural Change*, 36: 134-153
 - ▶ Mahapatra, K. & Kent, S. (2005). Tropical deforestation: A multinomial logistic model and some country-specific policy prescriptions. *Forest Policy and Economics*.
 - ▶ Peng, C.J. & Nichols, R.N. (2003). Using multinomial logistic models to predict adolescent behavioral risk. *Journal of Modern Applied Statistical Methods*. 2(1): 1-13.
-

