

การประมาณค่าและเปรียบเทียบผลกระทบ ด้วยคำสั่ง margins และ contrast ในโปรแกรม STATA

รศ.ดร.พงษ์เดช สารการ

สาขาวิทยาการระบาดและชีวสถิติ คณะสาธารณสุขศาสตร์ มหาวิทยาลัยขอนแก่น
กลุ่มวิจัยวิทยาการระบาดและป้องกันโรคมาเร็งในภูมิภาคอาเซียน (ACEP)

E-mail : spongde@kku.ac.th

Website : pongdechonline.com

1.1 บทนำ

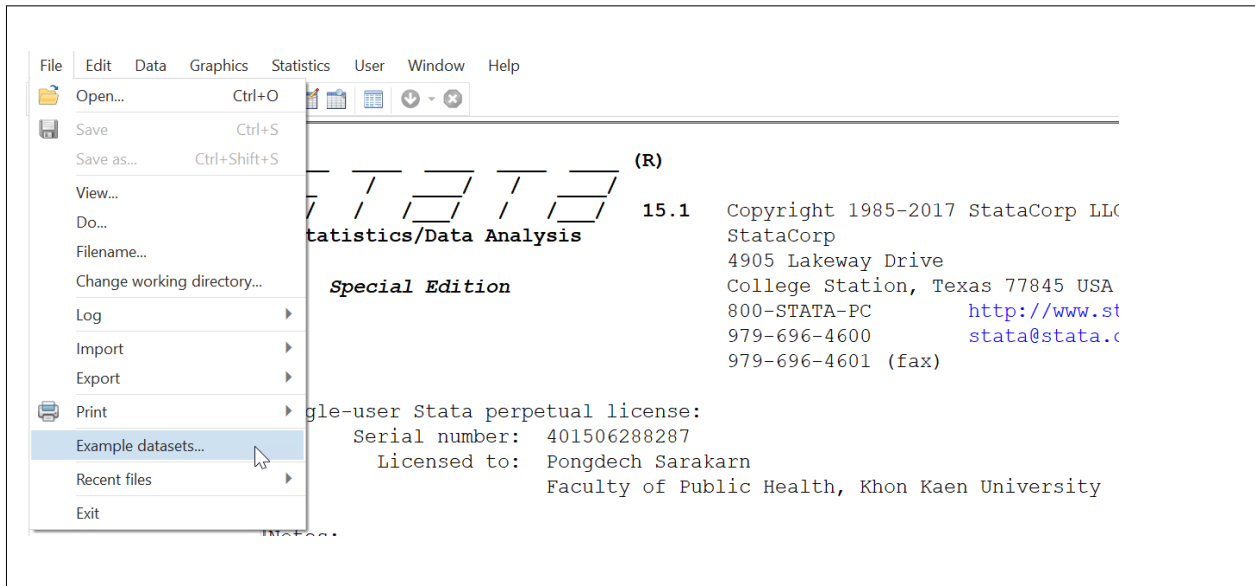
ในการวิเคราะห์สมการถดถอย เมื่อตัวแปรผลลัพธ์มีค่าต่อเนื่อง การประมาณค่าผลกระทบ ด้วยค่าประมาณที่ได้จากผลลัพธ์ หรือ ด้วยค่าสัมประสิทธิ์ถดถอย (regression coefficients) นักวิจัยสามารถนำไปใช้และแปลผลได้อย่างตรงไปตรงมา เช่น “เมื่อ X มีค่าเปลี่ยนแปลงไปหนึ่งหน่วย ค่า Y ที่ได้ ก็จะเปลี่ยนแปลงไปเท่ากับค่าสัมประสิทธิ์ถดถอยของ X ” เป็นต้น ส่วนการทำนายค่าตัวแปรตาม นักวิจัยก็สามารถทำได้ ทั้งในกรณีนำตัวแปรร่วม (covariate) มาพิจารณาแบบหนึ่งตัว หรือ มากกว่าหนึ่งตัวได้อย่างตรงไปตรงมาเช่นเดียวกัน แต่ในบางกรณี หากนักวิจัยสนใจผลกระทบที่เกิดขึ้นของปฏิสัมพันธ์ (interaction effects) หรือ กรณีตัวแปรผลลัพธ์ที่นำมาใช้ทำการวิเคราะห์สมการถดถอยในรูปแบบแจกแจง หรือ ตัวแปรกลุ่ม (categorical variables) หรือ กรณีรูปแบบฟังก์ชันไม่ใช่เชิงเส้น (nonlinear functional forms) ผลลัพธ์ที่ได้ อาจมีข้อจำกัดในการอธิบายเกี่ยวกับการประมาณและทำนายค่าผลกระทบที่เกิดขึ้น ทั้งจากตัวแปรร่วมหนึ่งตัว หรือ มากกว่าอย่างตรงไปตรงมาได้ เพราะค่าสัมประสิทธิ์ถดถอยของสมการถดถอยลอจิสติก อยู่ในรูป “log odd” เป็นต้น นอกจากนี้ยังรวมถึงการประมาณและทำนายค่าสำหรับวิธีการทางสถิติที่อยู่ภายใต้ตัวแบบเชิงเส้นทั่วไป (general linear model, GLM) อาทิ การวิเคราะห์ความแปรปรวน (ANOVA) หรือ การวิเคราะห์ความแปรปรวนร่วม (ANCOVA) เป็นต้น ดังนั้นจากข้อจำกัดที่เกิดขึ้นดังกล่าว โปรแกรม STATA จึงได้พัฒนาคำสั่ง **margins** ขึ้นมา เพื่อช่วยให้นักวิจัยสามารถนำมาใช้ประโยชน์ในการประมาณและทำนายค่าผลกระทบที่เกิดขึ้นภายใต้วิธีการทางสถิติต่างๆ ได้อย่างครอบคลุม โดยคำสั่ง **margins** ถือเป็นคำสั่งที่ถูกนำมาใช้ เพื่อให้โปรแกรม STATA คำนวณและแสดงผลลัพธ์เกี่ยวกับค่า margin ซึ่งเป็นค่าประมาณที่ถูกคำนวณได้โดยตรงจากสมการทำนายค่า ภายใต้ตัวแบบเชิงเส้น (linear model) หรือ ตัวแบบที่ไม่ใช่เชิงเส้น (non-linear model) ที่นักวิจัยเลือกนำมาใช้ในการวิเคราะห์ข้อมูล ซึ่งจากคำสั่งนี้ นักวิจัยยังสามารถกำหนดให้มีการนำตัวแปรร่วม (covariate variables) ภายในตัวแบบดังกล่าว มาพิจารณา โดยมีความเป็นไปได้ ครอบคลุมทั้งการกำหนดให้ตัวแปรร่วมไม่มีการเปลี่ยนแปลง หรือ คงที่ทุกตัวแปร (conditional margin) และสามารถกำหนดให้ตัวแปรร่วมมีการเปลี่ยนแปลงค่าได้บางส่วน (predictive margin) อีกด้วย

ในทางปฏิบัติ เพื่อให้ให้นักวิจัยได้เรียนรู้และฝึกการใช้วิธีการประมาณและทำนายค่าผลกระทบได้อย่างถูกต้องจากคำสั่ง **margins** ผู้เขียนจึงได้ใช้นำชุดข้อมูลตัวอย่าง (example datasets) ที่มีอยู่ในโปรแกรม STATA มาใช้ประกอบเป็นกรณีศึกษาในตัวอย่างการวิเคราะห์ข้อมูลและคำนวณค่าประมาณต่างๆ ที่เกี่ยวข้อง พร้อมนำเสนอแนวทางการแปลผลและการนำผลลัพธ์ที่ได้ไปใช้ประโยชน์เพื่อตอบคำถามวิจัยที่เกี่ยวข้อง โดยมีแนวทางการใช้งานชุดข้อมูลดังกล่าว ดังนี้

คำสั่งเรียกใช้ชุดข้อมูล (dataset) ในโปรแกรม STATA

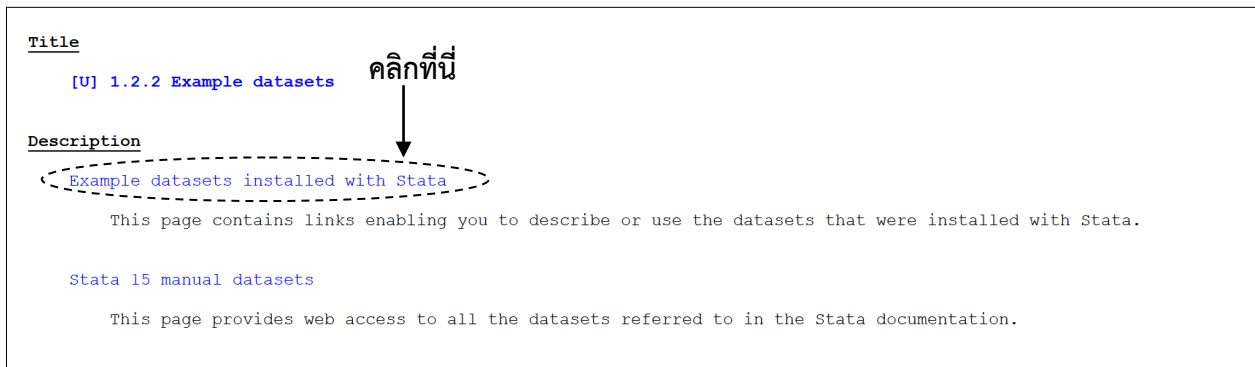
>>> กรณีไม่ทราบชื่อชุดข้อมูล

- ①. เลื่อนเมาส์ไปที่รายการหลัก File แล้วไปคลิกที่รายการย่อย *Example datasets...*
 ดังแผนภาพที่ 1.1



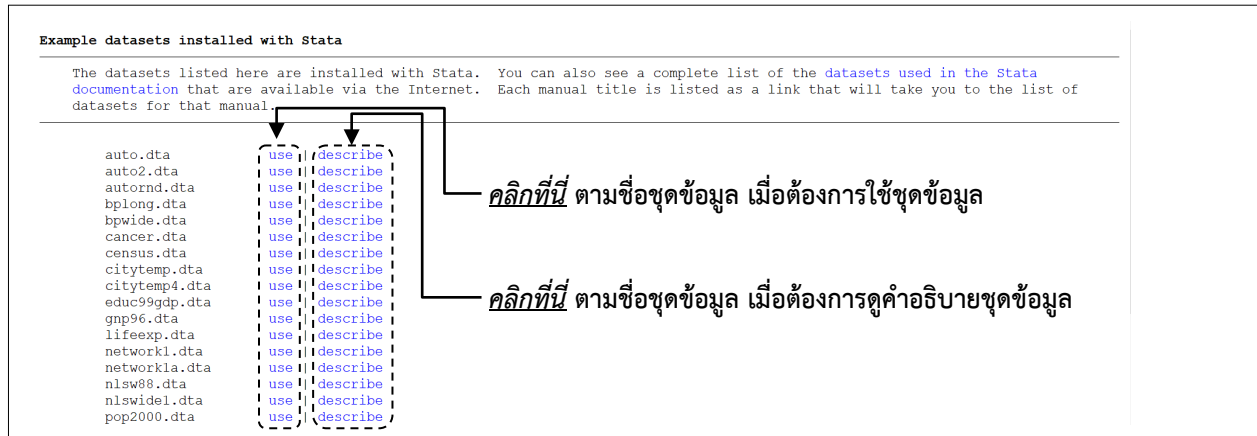
แผนภาพที่ 1.1 แสดงการเลือกรายการหลัก File และรายการย่อย *Example datasets...*

- ②. คลิกเมาส์ที่ *Example datasets installed with Stata* ดังแผนภาพที่ 1.2



แผนภาพที่ 1.2 แสดงการเลือกรายการหลัก File และรายการย่อย *Example datasets...*

๓. จากนั้นจะปรากฏรายการชุดข้อมูลให้นักวิจัยเลือก ซึ่งการเลือกชุดข้อมูล นักวิจัยสามารถทำได้ ด้วยการเลื่อนเมาส์ไปคลิกที่ *use* หรือหากต้องการดูคำอธิบายก็สามารถทำได้ ด้วยการเลื่อนเมาส์ไปคลิกที่ *describe* ในรายการชุดข้อมูลที่นักวิจัยต้องการเลือก ดังแผนภาพที่ 1.3



แผนภาพที่ 1.3 แสดงการเลือกชุดข้อมูลและคำอธิบายชุดข้อมูลที่ต้องการ ด้วยคำสั่ง *use* และ *describe*

>>> กรณีทราบชื่อชุดข้อมูล

รูปแบบคำสั่งการเลือกชุดข้อมูลในโปรแกรม STATA

```
sysuse dataset_name
```

sysuse	คำสั่งในการเลือกชุดข้อมูลในโปรแกรม STATA
dataset_name	ชื่อชุดข้อมูลในโปรแกรม STATA ที่นักวิจัยต้องการนำมาใช้งาน (ซึ่งจะใส่ .dta ตามหลังชื่อไฟล์ หรือไม่ใส่ ก็ได้) (ตัวเลือกอื่น ๆ สามารถศึกษาได้จากคำสั่ง help sysuse)

กรณีตัวอย่างที่ 1.1

กำหนดให้นักวิจัย เลือกใช้ชุดข้อมูล (dataset) ที่มีอยู่ในโปรแกรม STATA ด้วยคำสั่ง **sysuse** ภายใต้ชื่อ **bpwide.dta** โดยมีคำอธิบายตัวแปร ดังนี้

คำอธิบายตัวแปร

ชื่อตัวแปร	คำอธิบายชื่อตัวแปร	ค่าที่เป็นไปได้
patient	รหัสผู้ป่วย	ตามระบุ
sex	เพศ	0= ชาย , 1= หญิง
agegrp	ช่วงอายุ (ปี)	1= 30-45 ปี 2= 46-59 ปี 3= 60 ปีขึ้นไป
bp_before	ระดับความดันเลือด systolic ก่อนทดลอง (mmHg)	ตามระบุ
bp_after	ระดับความดันเลือด systolic หลังทดลอง (mmHg)	ตามระบุ

ตัวอย่างรายละเอียดตัวแปรและข้อมูลในชุดข้อมูล **bpwide.dta**

<pre>. sysuse bpwide.dta (fictional blood-pressure data)</pre>						
	patient	sex	agegrp	bp_before	bp_after	
1	1	Male	30-45	143	153	
2	2	Male	30-45	163	170	
3	3	Male	30-45	153	168	
4	4	Male	30-45	153	142	
5	5	Male	30-45	146	141	
6	6	Male	30-45	150	147	
7	7	Male	30-45	148	133	
8	8	Male	30-45	153	141	
9	9	Male	30-45	153	131	
10	10	Male	30-45	158	125	
11	11	Male	30-45	149	164	
12	12	Male	30-45	173	159	
13	13	Male	30-45	165	135	
14	14	Male	30-45	145	159	
15	15	Male	30-45	143	153	
16	16	Male	30-45	152	126	
17	17	Male	30-45	141	162	
18	18	Male	30-45	176	134	
19	19	Male	30-45	143	136	
20	20	Male	30-45	162	150	
21	21	Male	46-59	149	168	
22	22	Male	46-59	156	155	
23	23	Male	46-59	153	136	

แผนภาพที่ 1.4 แสดงตัวอย่างรายละเอียดตัวแปรและข้อมูลในชุดข้อมูล **bpwide.dta**

1.2 การประมาณและทำนายค่าผลกระทบ เมื่อกำหนดให้ตัวแปรร่วมทุกตัวคงที่ (Conditional margin)

จากกรณีตัวอย่างที่ 1.1 หากนักวิจัยทำการวิเคราะห์ข้อมูล เพื่อตอบคำถามวิจัยเกี่ยวกับ “เพศชาย หรือ หญิง มีความสัมพันธ์ หรือ มีผลต่อการเปลี่ยนแปลงระดับความดันเลือด systolic หลังการทดลอง หรือไม่ ?” นักวิจัยสามารถเลือกใช้วิธีสมการถดถอยแบบง่ายมาใช้ในการวิเคราะห์ข้อมูล โดยมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.5

. regress bp_after i.sex						
Source	SS	df	MS	Number of obs	=	120
Model	2075.00833	1	2075.00833	F(1, 118)	=	11.21
Residual	21844.5833	118	185.123588	Prob > F	=	0.0011
				R-squared	=	0.0867
				Adj R-squared	=	0.0790
Total	23919.5917	119	201.004972	Root MSE	=	13.606

bp_after	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
sex						
Female	-8.316667	2.484107	-3.35	0.001	-13.23587	-3.397459
_cons	155.5167	1.756529	88.54	0.000	152.0383	158.9951

แผนภาพที่ 1.5 แสดงผลลัพธ์ที่ได้จากการวิเคราะห์สมการถดถอยแบบง่าย เมื่อตัวแปรร่วมมี 1 ตัวแปร

จากแผนภาพที่ 1.5 นักวิจัยสามารถนำผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลดังกล่าว มาสรุปเป็นสมการทำนาย เพื่อตอบโจทย์คำถามวิจัยที่เกี่ยวข้อง ได้ดังนี้

$$[bp_after] = 155.51 - 8.32[sex_{Female}]$$

จากสมการทำนายที่ได้ดังกล่าว นักวิจัยสามารถนำมาพิจารณาผลกระทบที่เกิดขึ้นต่อตัวแปรผลลัพธ์ (bp_after) โดยจำแนกตามระดับของตัวแปรร่วมที่สนใจ ได้ดังนี้

$$\text{กรณีเพศชาย (sex=0)} \quad [bp_after] = 155.51 - 8.32[0] = 155.51$$

$$\text{กรณีเพศหญิง (sex=1)} \quad [bp_after] = 155.51 - 8.32[1] = 147.19$$

ซึ่งค่าประมาณที่คำนวณได้จากสมการถดถอยดังกล่าวข้างต้น นักวิจัยสามารถนำคำสั่ง **margins** ในโปรแกรม STATA มาใช้ในการคำนวณค่าที่เกี่ยวข้องดังกล่าวได้ โดยผลลัพธ์ หรือ คำตอบที่ได้รับ ยังคงให้ผลเช่นเดียวกัน ดังแผนภาพที่ 1.6

```
. margins i.sex
```

Adjusted predictions

Number of obs

=

120

Model VCE : OLS

Expression : Linear prediction, predict()

		Delta-method				
	Margin	Std. Err.	t	P> t	[95% Conf. Interval]	
sex						
Male	155.5167	1.756529	88.54	0.000	152.0383	158.9951
Female	147.2	1.756529	83.80	0.000	143.7216	150.6784

แผนภาพที่ 1.6 แสดงผลลัพธ์ที่ได้จากการประมาณและทำนายค่าผลกระทบด้วยคำสั่ง **margins**

จากแผนภาพที่ 1.6 พบว่า ผลลัพธ์ที่ได้จากการใช้คำสั่ง **margins** สามารถให้ค่าประมาณที่ได้จากการคำนวณเช่นเดียวกับที่นักวิจัยคำนวณเองด้วยมือ อย่างไรก็ตาม ผลลัพธ์ที่ได้จากการคำนวณด้วยคำสั่ง **margins** ของโปรแกรม STATA ยังให้ค่าสถิติที่เกี่ยวข้องได้ครอบคลุมมากกว่าโดยเฉพาะค่าช่วงเชื่อมั่น (95% CI) ซึ่งนักวิจัยสามารถนำค่าดังกล่าว ไปใช้ประโยชน์ในการประมาณค่าพารามิเตอร์ระดับประชากรได้ ขณะเดียวกันเมื่อจำนวนตัวแปรรวม ที่นำมาวิเคราะห์ในตัวแบบดังกล่าวมีจำนวนเพิ่มมากขึ้น การคำนวณเองด้วยมือตามแนวทางข้างต้นอาจไม่เหมาะสม เพราะมีขั้นตอนค่อนข้างยุ่งยากและซับซ้อน ดังนั้นการคำนวณด้วยคำสั่ง **margins** จึงมีความสะดวกและครอบคลุมค่าผลลัพธ์ที่ได้มากกว่า ดังกรณีตัวอย่างที่ 1.1 หากนักวิจัยทำการวิเคราะห์ข้อมูลเพื่อตอบคำถามวิจัยเกี่ยวกับ “**เพศและช่วงอายุ มีความสัมพันธ์ หรือ มีผลต่อการเปลี่ยนแปลงระดับความดันเลือด systolic หลังการทดลอง หรือไม่ ?**” นักวิจัยสามารถเลือกใช้วิธีสมการถดถอยแบบง่ายมาใช้ในการวิเคราะห์ข้อมูล โดยมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.7

. regress bp_after i.sex i.agegrp						
Source	SS	df	MS	Number of obs	= 120	
Model	6387.875	3	2129.29167	F(3, 116)	= 14.09	
Residual	17531.7167	116	151.135489	Prob > F	= 0.0000	
Total	23919.5917	119	201.004972	R-squared	= 0.2671	
				Adj R-squared	= 0.2481	
				Root MSE	= 12.294	
bp_after	Coef.	Std. Err.	t	P> t 	[95% Conf. Interval]	
sex						
Female	-8.316667	2.244515	-3.71	0.000	-12.76221	-3.871121
agegrp						
46-59	6.45	2.748959	2.35	0.021	1.005341	11.89466
60+	14.65	2.748959	5.33	0.000	9.205341	20.09466
_cons	148.4833	2.244515	66.15	0.000	144.0378	152.9289

แผนภาพที่ 1.7 แสดงผลลัพธ์ที่ได้จากการวิเคราะห์สมการถดถอยแบบง่าย เมื่อตัวแปรรวมมี 2 ตัวแปร

$$[\text{bp_after}] = 155.51 - 8.32[\text{sex}_{\text{Female}}] + 6.45[\text{agegrp}_{46-49}] + 14.65[\text{agegrp}_{60+}]$$

ซึ่งค่าประมาณที่คำนวณได้จากสมการถดถอยดังกล่าวข้างต้น นักวิจัยสามารถนำคำสั่ง **margins** ในโปรแกรม STATA มาใช้ในการคำนวณค่าที่เกี่ยวข้องดังกล่าวได้ โดยมีผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.8

```
. margins i.sex i.agegrp
```

Predictive margins Number of obs = 120
Model VCE : OLS

Expression : Linear prediction, predict()

```
-----+-----
```

		Delta-method				
	Margin	Std. Err.	t	P> t	[95% Conf. Interval]	
sex						
Male	155.5167	1.587112	97.99	0.000	152.3732	158.6601
Female	147.2	1.587112	92.75	0.000	144.0565	150.3435
agegrp						
30-45	144.325	1.943807	74.25	0.000	140.475	148.175
46-59	150.775	1.943807	77.57	0.000	146.925	154.625
60+	158.975	1.943807	81.79	0.000	155.125	162.825

```
-----+-----
```

แผนภาพที่ 1.8 แสดงผลลัพธ์ที่ได้จากการประมาณและทำนายค่าผลกระทบของตัวแปรร่วม sex และ agegrp

นอกจากการประมาณและทำนายค่าผลกระทบที่เกิดขึ้นจากตัวแปรร่วมหนึ่งตัว หรือ มากกว่า ภายหลังคำสั่ง **margins** โปรแกรม STATA ได้พัฒนาคำสั่ง **marginsplot** เพื่อนำผลลัพธ์ ซึ่งถูกคำนวณด้วยคำสั่ง **margins** ไปก่อนหน้านี้ มาแสดงผลเป็นกราฟเส้น พร้อมช่วงเชื่อมั่น 95%CI ประกอบ ดังแผนภาพที่ 1.9

```
. margins i.agegrp
```

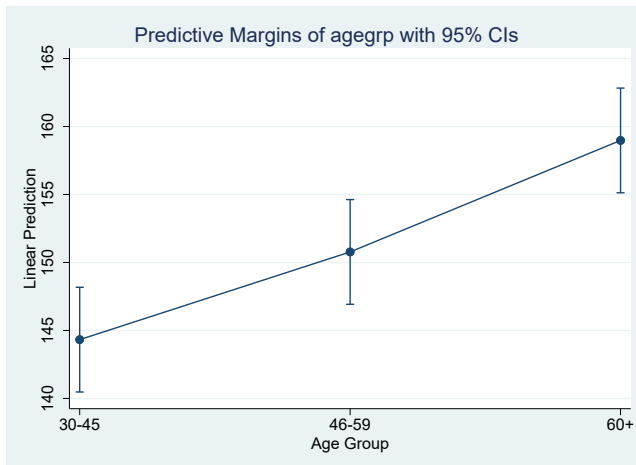
```
Predictive margins                                Number of obs    =          120
Model VCE      : OLS

Expression     : Linear prediction, predict()
```

		Delta-method				[95% Conf. Interval]	
		Margin	Std. Err.	t	P> t		
agegrp							
30-45		144.325	1.943807	74.25	0.000	140.475	148.175
46-59		150.775	1.943807	77.57	0.000	146.925	154.625
60+		158.975	1.943807	81.79	0.000	155.125	162.825

```
. marginsplot
```

```
Variables that uniquely identify margins: agegrp
```



แผนภาพที่ 1.9 แสดงผลลัพธ์ที่ได้จากการใช้คำสั่ง **margins** และ **marginsplot**

นอกจากนี้ นักวิจัยยังสามารถประเมินผลกระทบร่วม หรือ การมีปฏิสัมพันธ์กัน (interaction) ระหว่างตัวแปรสองตัว หรือ มากกว่าได้ โดยจากกรณีตัวอย่างที่ 1.1 หากนักวิจัยสนใจผลกระทบร่วมที่เกิดขึ้นระหว่างเพศ (sex) กับช่วงอายุ (agegrp) สามารถคำนวณและแสดงผลลัพธ์ได้ ด้วยการใช้คำสั่ง **margins** และคำสั่ง **marginsplot** รวมถึงการแปลความหมายผลกระทบที่เกิดขึ้น ดังแผนภาพที่ 1.10

```
. quietly regress bp_after i.sex i.agegrp
```

เป็นคำสั่งที่ไม่ให้แสดงผลลัพธ์ของการวิเคราะห์
แต่ยังคงมีการบันทึกไว้ในหน่วยความจำแทน

```
. margins sex#agegrp
```

Adjusted predictions

Number of obs = 120

Model VCE : OLS

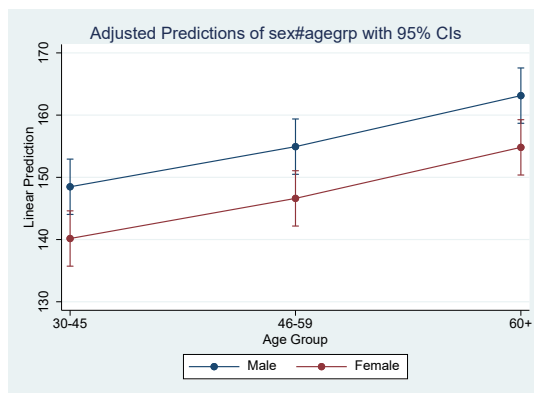
Expression : Linear prediction, predict()

	Delta-method					
	Margin	Std. Err.	t	P> t	[95% Conf. Interval]	
sex#agegrp						
Male#30-45	148.4833	2.244515	66.15	0.000	144.0378	152.9289
Male#46-59	154.9333	2.244515	69.03	0.000	150.4878	159.3789
Male#60+	163.1333	2.244515	72.68	0.000	158.6878	167.5789
Female#30-45	140.1667	2.244515	62.45	0.000	135.7211	144.6122
Female#46-59	146.6167	2.244515	65.32	0.000	142.1711	151.0622
Female#60+	154.8167	2.244515	68.98	0.000	150.3711	159.2622

ผู้ชายที่มีอายุ 60 ปีขึ้นไป มีระดับความดันเลือด systolic ค่อนข้างสูงกว่าทุกกลุ่ม (โดยมีค่าประมาณเท่ากับ 163.13, 95%CI=158.69-167.58)

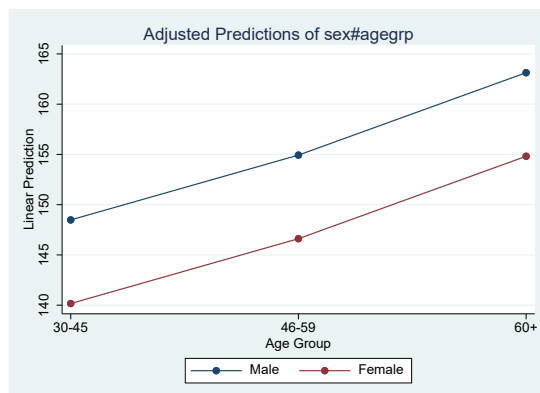
```
. marginsplot, x(agegrp)
```

Variables that uniquely identify margins: sex agegrp



```
. marginsplot, x(agegrp) noci
```

Variables that uniquely identify margins: sex agegrp



แผนภาพที่ 1.10 แสดงผลลัพธ์ที่ได้จากการใช้คำสั่ง **margins** และ **marginsplot** กรณีผลกระทบร่วม

ดังนั้นการประมาณและทำนายค่าผลกระทบของสมการถดถอย ภายใต้ตัวแบบเชิงเส้นและตัวแบบที่ไม่ใช่เชิงเส้น (linear and non-linear model) กรณีตัวแปรร่วมทุกตัวคงที่ หรือ เรียกว่า conditional margin นักวิจัยสามารถใช้คำสั่ง **margins** เพื่อคำนวณค่าผลกระทบที่เกิดขึ้นในแต่ละตัวแปรร่วมที่สนใจ รวมถึงผลกระทบที่เกิดขึ้นจากการมีปฏิสัมพันธ์กันและสามารถใช้คำสั่ง **marginsplot** เพื่อแสดงผลเป็นกราฟเส้น พร้อมช่วงเชื่อมั่น 95%CI ประกอบ และที่สำคัญนักวิจัยต้องคำนึงถึงเสมอว่า คำสั่ง **margins** และคำสั่ง **marginsplot** ถือเป็น เป็นคำสั่งที่ถูกนำมาใช้หลังการประมาณค่า หรือ postestimation นั่นคือการใช้คำสั่ง **margins** ต้องใช้ภายหลัง จากที่ได้มีการใช้คำสั่งในการวิเคราะห์ข้อมูล ตามตัวแบบถดถอยที่นักวิจัยเลือกก่อนเสมอ ขณะเดียวกันการใช้คำสั่ง **marginsplot** ก็ต้องถูกใช้หลังจากเสร็จสิ้นการใช้คำสั่ง **margins** ผ่านไปแล้วเช่นเดียวกัน

1.3 การประมาณและทำนายค่าผลกระทบ เมื่อกำหนดให้ตัวแปรร่วมเปลี่ยนแปลงค่าได้ (Predictive margin)

ในทางปฏิบัติสำหรับการประมาณและทำนายค่าผลกระทบ กรณีตัวแปรร่วมสามารถเปลี่ยนแปลงค่าได้ ประกอบด้วย

- 1.3.1 การประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าเฉลี่ย
(marginal effects at the mean; MEM)
- 1.3.2 การประมาณและทำนายค่าผลกระทบโดยเฉลี่ย
(average marginal effects; AME)
- 1.3.3 การประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าระบุที่เป็นตัวแทน
(marginal effects at representative values; MER)

กรณีตัวอย่างที่ 1.2

กำหนดให้นักวิจัย เลือกใช้ชุดข้อมูล (dataset) ที่มีอยู่ในโปรแกรม STATA ด้วยคำสั่ง **sysuse** ภายใต้ชื่อ **cancer.dta** โดยมีคำอธิบายตัวแปร ดังนี้

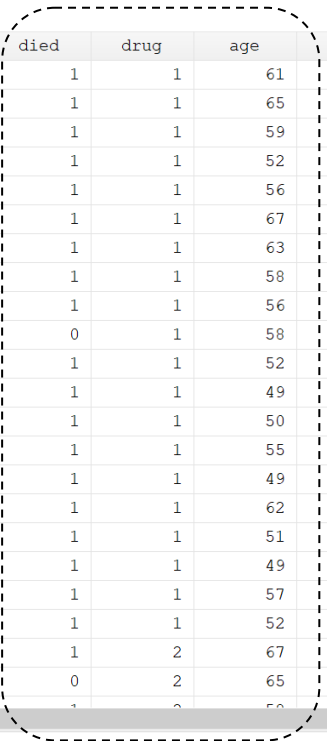
คำอธิบายตัวแปร*

ชื่อตัวแปร	คำอธิบายชื่อตัวแปร	ค่าที่เป็นไปได้
died	การเสียชีวิตของผู้ป่วยมะเร็ง	0= มีชีวิต , 1= เสียชีวิต
drug	ชนิดของยา	1= ยาหลอก (placebo) 2= ยาชนิด A 3= ยาชนิด B
age	อายุ (ปี)	ตามระบุ

*อธิบายเฉพาะตัวแปรที่ถูกนำมาใช้งานเท่านั้น

ตัวอย่างรายละเอียดตัวแปรและข้อมูลในชุดข้อมูล `cancer.dta`

`. sysuse cancer.dta`
(Patient Survival in Drug Trial)



	studytime	died	drug	age	_st	_d	_t	_t0
1	1	1	1	61	1	1	1	0
2	1	1	1	65	1	1	1	0
3	2	1	1	59	1	1	2	0
4	3	1	1	52	1	1	3	0
5	4	1	1	56	1	1	4	0
6	4	1	1	67	1	1	4	0
7	5	1	1	63	1	1	5	0
8	5	1	1	58	1	1	5	0
9	8	1	1	56	1	1	8	0
10	8	0	1	58	1	0	8	0
11	8	1	1	52	1	1	8	0
12	8	1	1	49	1	1	8	0
13	11	1	1	50	1	1	11	0
14	11	1	1	55	1	1	11	0
15	12	1	1	49	1	1	12	0
16	12	1	1	62	1	1	12	0
17	15	1	1	51	1	1	15	0
18	17	1	1	49	1	1	17	0
19	22	1	1	57	1	1	22	0
20	23	1	1	52	1	1	23	0
21	6	1	2	67	1	1	6	0
22	6	0	2	65	1	0	6	0
23	7	1	2	50	1	1	7	0

แผนภาพที่ 1.11 แสดงตัวอย่างรายละเอียดตัวแปรและข้อมูลในชุดข้อมูล `cancer.dta`

1.3.1 การประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าเฉลี่ย (marginal effects at the mean; MEM)

การประมาณค่าผลกระทบ ณ ตำแหน่งค่าเฉลี่ย (marginal effects at the mean; MEM) เป็นวิธีการพิจารณาค่าผลกระทบภายใต้แนวคิดดังนี้

“กำหนดให้ตัวแปรร่วมทุกตัวมีค่าคงที่ เท่ากับ ค่ากลาง หรือ ค่าเฉลี่ยของตัวมันเอง”

จากกรณีตัวอย่างที่ 1.2 หากนักวิจัยทำการวิเคราะห์ข้อมูล เพื่อตอบคำถามวิจัยเกี่ยวกับ “ชนิดของยาและอายุ มีความสัมพันธ์ หรือ มีผลต่อการเสียชีวิตของผู้ป่วยโรคมะเร็ง หรือไม่ ?” เนื่องจากตัวแปรผลลัพธ์ มีค่าแบบแจกนับ 2 ค่า (binary values) ดังนั้นนักวิจัยจึงสามารถเลือกใช้วิธีสมการถดถอยแบบลอจิสติก (logistic regression) มาทำการวิเคราะห์ข้อมูลได้ โดยมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.12

. logit died i.drug age

```

Iteration 0: log likelihood = -31.199418
Iteration 1: log likelihood = -22.259892
Iteration 2: log likelihood = -21.539075
Iteration 3: log likelihood = -21.525583
Iteration 4: log likelihood = -21.525537
Iteration 5: log likelihood = -21.525537

```

Logistic regression

Number of obs = 48
LR chi2(3) = 19.35
Prob > chi2 = 0.0002
Pseudo R2 = 0.3101

Log likelihood = -21.525537

died	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	

drug						
2	-3.550783	1.222919	-2.90	0.004	-5.94766	-1.153905
3	-3.249504	1.180797	-2.75	0.006	-5.563824	-.935184
age	.1203223	.0718246	1.68	0.094	-.0204513	.2610959
_cons	-3.625338	3.966999	-0.91	0.361	-11.40051	4.149838

. margins drug

Predictive margins
Model VCE : OIM

โอกาส หรือความน่าจะเป็นที่ผู้รับยาชนิดที่ 1
จะเสียชีวิต เท่ากับ 0.948 (95%CI= 0.85-1.00)

Number of obs = 48

Expression : Pr(died),

predict()

	Margin	Delta-method Std. Err.	z	P> z	[95% Conf. Interval]	

drug						
1	.9483949	.0497326	19.07	0.000	.8509208	1.045869
2	.3977432	.1262881	3.15	0.002	.150223	.6452634
3	.4642472	.1263103	3.68	0.000	.2166836	.7118107

แผนภาพที่ 1.12 แสดงผลลัพธ์ที่ได้จากการวิเคราะห์ข้อมูลและประมาณค่าด้วยคำสั่ง **margins** กรณีตัวอย่างที่ 4.3

จากแผนภาพที่ 1.12 พบว่า การประมาณและทำนายค่าผลกระทบ ถูกพิจารณาแบบตัวแปรร่วมคงที่ (conditional margin) ซึ่งในบางกรณีผลกระทบที่ถูกพิจารณา อาจมีความลำเอียงจากความเบี่ยงเบนของข้อมูลดิบ หรือ สัดส่วนการกระจายตัวที่ไม่เท่ากันจากเหตุการณ์ที่สนใจในแต่ละระดับของตัวแปรร่วม (ซึ่งในกรณีนี้ ได้แก่ การเสียชีวิต) ดังนั้นการนำวิธีการประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าเฉลี่ย จึงถูกนำมาใช้ ด้วยคำสั่ง **margins** เดิม แต่เพิ่มตัวเลือก (option) ใหม่เข้ามา ได้แก่ **atmeans** ซึ่งมีผลลัพธ์ที่ได้แตกต่างจากกรณีตัวแปรร่วมคงที่ ดังแผนภาพที่ 1.13

```
. margins drug, atmeans
```

Adjusted predictions
Model VCE : OIM

Expression : Pr(died), predict()
at : 1.drug = .4166667 (mean)
2.drug = .2916667 (mean)
3.drug = .2916667 (mean)
age = 55.875 (mean)

ในแต่ละระดับของตัวแปรรวม ถูกกำหนดให้คงที่เท่ากับค่ากลาง หรือค่าเฉลี่ยของตัวเอง

	Margin	Std. Err.	z	P> z	[95% Conf. Interval]
drug					
1	.9567965	.0434103	22.04	0.000	.8717139 1.041879
2	.3886211	.1404264	2.77	0.006	.1133905 .6638517
3	.4621143	.1398518	3.30	0.001	.1880099 .7362188

แผนภาพที่ 1.13 แสดงผลลัพธ์ที่ได้ด้วยคำสั่ง **margins** ตามด้วยตัวเลือก **atmeans** ในกรณีตัวอย่างที่ 1.2

1.3.2 การประมาณและทำนายค่าผลกระทบโดยเฉลี่ย (average marginal effects; AME)

การประมาณและทำนายค่าผลกระทบโดยเฉลี่ย (average marginal effects; AME) เป็นวิธีการพิจารณาค่าผลกระทบภายใต้แนวคิดดังนี้

“ค่าผลกระทบที่ได้ ถูกพิจารณาจากสัดส่วนระดับที่สนใจ (dy) เทียบกับระดับพื้นฐานระดับเริ่มต้น”

จากกรณีตัวอย่างที่ 1.2 ภายใต้การวิเคราะห์หัตถถอยลอจิสติกเช่นเดิม นักวิจัยสามารถประมาณและทำนายค่าผลกระทบโดยเฉลี่ยได้ โดยมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.14

```
. quietly logit died i.drug age
```

```
. margins, dydx(drug)
```

Average marginal effects
Model VCE : OIM

Expression : Pr(died), predict()
dy/dx w.r.t. : 2.drug 3.drug

โดยเฉลี่ยแล้ว ผู้ป่วยที่ได้รับยาชนิด A มีโอกาส หรือความน่าจะเป็นที่จะเสียชีวิตด้วยโรคมะเร็ง น้อยกว่า ผู้ป่วยที่รับยาหลอก เท่ากับ 0.55 (95%CI= 0.28-0.81)

	dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]
drug_1					
2	-.5506517	.1355682	-4.06	0.000	-.8163604 -.284943
3	-.4841477	.1359077	-3.56	0.000	-.750522 -.2177734

Note: dy/dx for factor levels is the discrete change from the base level.

แผนภาพที่ 1.14 แสดงผลลัพธ์ที่ได้ด้วยคำสั่ง **margins** ตามด้วยตัวเลือก **dydx** ในกรณีตัวอย่างที่ 1.2

จากแผนภาพที่ 1.14 พบว่า การประมาณและทำนายค่าผลกระทบ ถูกพิจารณาแบบสัดส่วน โดยเทียบกับระดับแรกสุดและถือเป็นจุดอ้างอิง ดังเช่นกรณีข้างต้น ค่าประมาณผลกระทบที่ได้ (dy/dx) ของ drug=2 หมายถึง “ผลกระทบที่เกิดขึ้นจากสัดส่วนของผู้ป่วยที่ได้รับยาชนิด A เทียบกับผู้ป่วยที่ได้รับยาหลอก เท่ากับ -0.55 (เครื่องหมายลบ แสดงถึง ความเสี่ยงที่ลดน้อยลงกว่าตัวเทียบ)” ขณะที่ (dy/dx) ของ drug=3 หมายถึง “ผลกระทบที่เกิดขึ้นจากสัดส่วนของผู้ป่วยที่ได้รับยาชนิด B เทียบกับผู้ป่วยที่ได้รับยาหลอก เท่ากับ -0.48”

1.3.3 การประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าระบุที่เป็นตัวแทน (marginal effects at representative values; MER)

การประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าระบุที่เป็นตัวแทน (marginal effects at representative values; MER) เป็นวิธีการพิจารณาผลกระทบภายใต้แนวคิดดังนี้

“ค่าผลกระทบที่ได้ ถูกพิจารณาจากการกำหนดค่าตัวแปรร่วม ภายใต้ความสนใจของนักวิจัยและสามารถกำหนดได้ทั้งแบบค่าต่อเนื่องและแบบการระบุเป็นค่าเฉพาะ”

จากกรณีตัวอย่างที่ 1.2 ภายใต้การวิเคราะห์ถดถอยลอจิสติกเช่นเดิม นักวิจัยสามารถประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าระบุที่เป็นตัวแทน โดยมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.15

<pre>. quietly logit died i.drug age</pre>						
<pre>. margins drug, at (age=(50 (5) 70))</pre>						
Adjusted predictions Model VCE : OIM Expression : Pr(died), predict() 1. _at : age = 50 2. _at : age = 55 3. _at : age = 60 4. _at : age = 65 5. _at : age = 70						
Number of obs = 48						
Delta-method Margin Std. Err. z P> z [95% Conf. Interval]						

_at#drug						
1 1	.9161213	.0815467	11.23	0.000	.7562927	1.07595
1 2	.2386663	.1439049	1.66	0.097	-.0433822	.5207147
1 3	.2976054	.1381571	2.15	0.031	.0268224	.5683883
2 1	.9522291	.0473552	20.11	0.000	.8594146	1.045044
2 2	.3639191	.1404506	2.59	0.010	.088641	.6391972
2 3	.4360729	.1373382	3.18	0.001	.166895	.7052508
3 1	.9732472	.0296782	32.79	0.000	.9150789	1.031416
3 2	.5108026	.1513805	3.37	0.001	.2141024	.8075029
3 3	.5852822	.1619051	3.61	0.000	.2679541	.9026103
4 1	.985162	.0194626	50.62	0.000	.9470159	1.023308
4 2	.6558442	.1782697	3.68	0.000	.306442	1.005246
4 3	.7203316	.1841825	3.91	0.000	.3593406	1.081323
5 1	.9918149	.0128463	77.21	0.000	.9666366	1.016993
5 2	.776683	.1850711	4.20	0.000	.4139504	1.139416
5 3	.824584	.1762525	4.68	0.000	.4791355	1.170032

แผนภาพที่ 1.15 แสดงผลลัพธ์ที่ได้ด้วยคำสั่ง **margins** ตามด้วยตัวเลือก **at** ในกรณีตัวอย่างที่ 1.2

จากแผนภาพที่ 1.15 พบว่า การประมาณและทำนายค่าผลกระทบ ณ ตำแหน่งค่าระบุที่เป็นตัวแทน นักวิจัยสามารถตอบโจทย์ข้อสงสัย หรือ ข้อค้นพบบางอย่างที่ต้องการได้ ดังกรณีผลลัพธ์ข้างต้น เมื่อผู้ป่วยในกลุ่มที่ได้รับยาหลอกและมีอายุตั้งแต่ 50 ปี ถึง 70 ปี พบว่า มีโอกาสเสี่ยง หรือ ความน่าจะเป็นที่จะเสียชีวิต ด้วยโรคมะเร็งเพิ่มขึ้นทุก ๆ 5 ปี อย่างต่อเนื่อง นั่นคือ ตั้งแต่ 0.92, 0.95, 0.97, 0.98 และ 0.99 ตามลำดับ

อย่างไรก็ตาม นอกจากคำสั่ง **margins** จะสามารถนำมาใช้เพื่อตอบโจทย์ในแต่ละลักษณะ ดังที่ได้กล่าวไปข้างต้นแล้ว ยังสามารถนำมาใช้ประกอบการพิจารณา เพื่อตอบโจทย์ปัญหาวิจัย หรือ ข้อสงสัย ที่นักวิจัยสนใจได้ ดังแผนภาพที่ 1.16

```
. quietly logit died i.drug age
```

```
. margins, dydx(drug) atmeans
```

Conditional marginal effects				Number of obs	=	48
Model VCE : OIM						
Expression : Pr(died), predict()						
dy/dx w.r.t. : 2.drug 3.drug						
at	:	1.drug	=	.4166667	(mean)	
		2.drug	=	.2916667	(mean)	
		3.drug	=	.2916667	(mean)	
		age	=	55.875	(mean)	

		Delta-method				
		dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]

drug						
2		-.5681755	.1484252	-3.83	0.000	-.8590836 -.2772674
3		-.4946822	.1455505	-3.40	0.001	-.779956 -.2094084

Note: dy/dx for factor levels is the discrete change from the base level.

```
. margins, dydx(drug) at(age=(50 60 70))
```

Conditional marginal effects				Number of obs	=	48
Model VCE : OIM						
Expression : Pr(died), predict()						
dy/dx w.r.t. : 2.drug 3.drug						
1._at	:	age	=	50		
2._at	:	age	=	60		
3._at	:	age	=	70		

		Delta-method				
		dy/dx	Std. Err.	z	P> z	[95% Conf. Interval]

1.drug		(base outcome)				
--------	--	----------------	--	--	--	--

2.drug						
_at						
1		-.677455	.1541367	-4.40	0.000	-.9795573 -.3753527
2		-.4624446	.1506191	-3.07	0.002	-.7576526 -.1672365
3		-.2151319	.177181	-1.21	0.225	-.5624003 .1321365

3.drug						
_at						
1		-.6185159	.1517475	-4.08	0.000	-.9159356 -.3210962
2		-.387965	.1577428	-2.46	0.014	-.6971352 -.0787949
3		-.1672309	.1678788	-1.00	0.319	-.4962673 .1618055

Note: dy/dx for factor levels is the discrete change from the base level.

แผนภาพที่ 1.16 แสดงผลลัพธ์ที่ได้ด้วยคำสั่ง **margins** แบบผสมหลายรูปแบบตามประเด็นที่นักวิจัยสนใจ

ขณะเดียวกัน เพื่อให้การตอบโจทย์เกี่ยวกับการประมาณและทำนายค่าผลกระทบมีความครอบคลุม นอกจากการประมาณค่าเชิงขนาดที่ได้ ด้วยคำสั่ง **margins** แล้ว นักวิจัยยังสามารถเปรียบเทียบค่าประมาณที่ได้ในระหว่างกลุ่มที่สนใจได้ ด้วยคำสั่ง **contrast** อีกด้วย

1.4 การเปรียบเทียบค่าประมาณผลกระทบ ด้วยคำสั่ง **contrast** ใน STATA

ในทางปฏิบัติคำสั่ง **contrast** สามารถถูกนำมาใช้ได้ทั้งแบบคำสั่งหลัก หรือ แบบเป็นตัวเลือกเสริมภายใต้คำสั่งหลัก **margins** แต่อย่างไรก็ตามในบทความนี้จะขอกล่าวถึงคำสั่ง **contrast** เฉพาะในกรณีที่ถูกนำมาใช้ร่วมเป็นตัวเลือกเสริมภายใต้คำสั่งหลัก **margins** โดยมีประเด็นที่เกี่ยวข้อง ประกอบด้วย

- 1.4.1 การเปรียบเทียบแบบภาพรวม
- 1.4.2 การเปรียบเทียบแบบใช้ระดับพื้นฐานของตัวแปรร่วมเป็นจุดเทียบ
- 1.4.3 การเปรียบเทียบแบบใช้ระดับกลุ่มก่อนหน้า หรือถัดไปของตัวแปรร่วมเป็นจุดเทียบ
- 1.4.4 การเปรียบเทียบแบบรายคู่ที่เป็นไปได้ทั้งหมด
- 1.4.5 การเปรียบเทียบแบบใช้ระดับพื้นฐานของตัวแปรร่วมเป็นจุดเทียบ
- ณ ตำแหน่งที่นักวิจัยสนใจ
- 1.4.6 การเปรียบเทียบแบบรายคู่ระหว่างสองตัวแปรร่วม

1.4.1 การเปรียบเทียบแบบภาพรวม

การเปรียบเทียบแบบภาพรวม ด้วยคำสั่งหลัก **margins** และภายใต้ตัวเลือกคำสั่ง **contrast** เป็นวิธีการพิจารณาการเปรียบเทียบค่าผลกระทบภายใต้แนวคิดดังนี้

“ค่าผลกระทบที่ได้ทุกกลุ่ม ถูกนำมาเปรียบเทียบว่า มีความแตกต่างในภาพรวม หรือไม่ ?”

จากกรณีตัวอย่างที่ 1.2 โดยใช้ผลการวิเคราะห์ถดถอยลอจิสติกเช่นเดิม นักวิจัยสามารถเปรียบเทียบค่าผลกระทบภาพรวม โดยมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.17

. quietly logit died i.drug age			
. margins drug,contrast			
Contrasts of predictive margins			
Model VCE : OIM			
Expression : Pr(died), predict()			
	df	chi2	P>chi2
drug	2	26.12	0.0000

ค่าประมาณผลกระทบของยาทั้ง 3 ชนิด แตกต่างกันอย่างมีนัยสำคัญทางสถิติ ที่ระดับความผิดพลาดไม่เกิน 5% (p-value < 0.001)

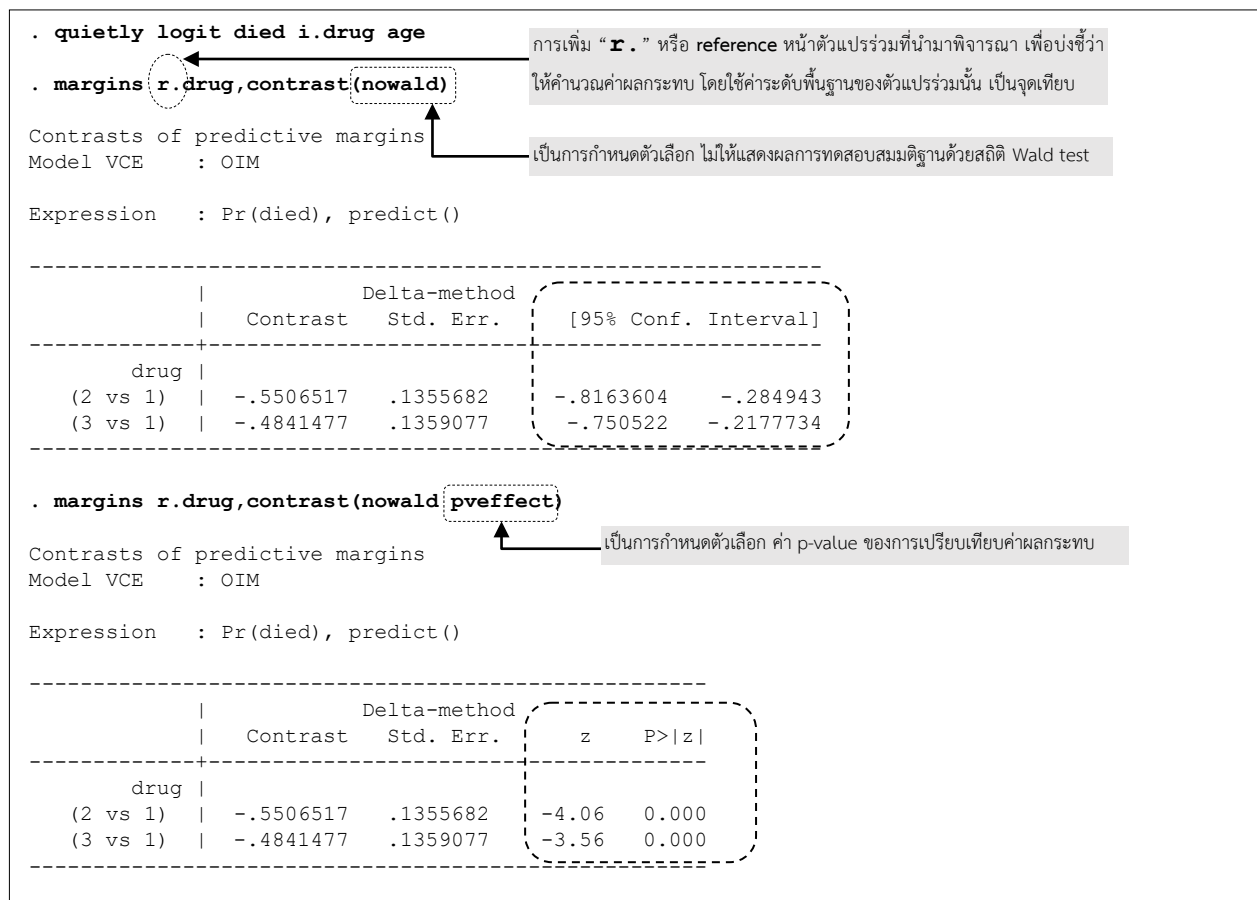
แผนภาพที่ 1.17 แสดงผลลัพธ์ที่ได้กรณีการเปรียบเทียบค่าประมาณผลกระทบแบบภาพรวม

1.4.2 การเปรียบเทียบแบบใช้ระดับพื้นฐานของตัวแปรร่วมเป็นจุดเทียบ

การเปรียบเทียบแบบใช้ระดับพื้นฐานของตัวแปรร่วมเป็นจุดเทียบ ด้วยคำสั่งหลัก **margins** และภายใต้ตัวเลือกคำสั่ง **contrast** เป็นวิธีการพิจารณาการเปรียบเทียบค่าผลกระทบภายใต้แนวคิดดังนี้

“ค่าผลกระทบที่ได้ในแต่ละระดับ ถูกนำมาเปรียบเทียบกับค่าระดับพื้นฐาน (base level) หรือ ค่าระดับแรกสุดของตัวแปรร่วมที่นำมาพิจารณา”

จากกรณีตัวอย่างที่ 1.2 โดยใช้ผลการวิเคราะห์ถดถอยลอจิสติกเช่นเดิม นักวิจัยสามารถเปรียบเทียบค่าผลกระทบแบบใช้ระดับพื้นฐานเป็นจุดเทียบ โดยการเพิ่ม “**r.**” หน้าตัวแปรร่วมที่ต้องการประมาณค่าและเปรียบเทียบผลกระทบ ซึ่งมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.18



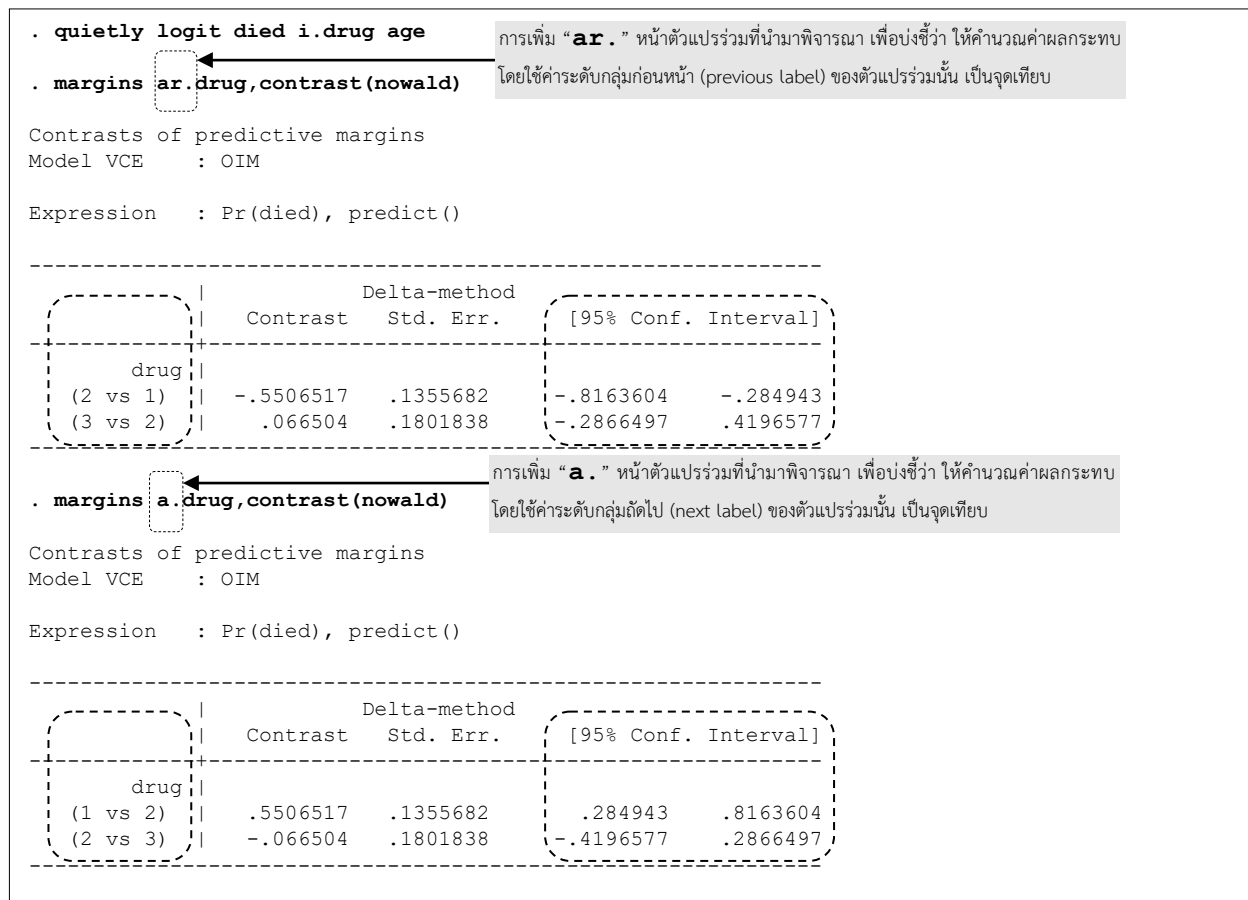
แผนภาพที่ 1.18 แสดงผลลัพธ์ที่ได้กรณีเปรียบเทียบค่าผลกระทบแบบใช้ระดับพื้นฐานเป็นจุดเทียบ

1.4.3 การเปรียบเทียบแบบใช้ระดับกลุ่มก่อนหน้า (previous label) หรือ ถัดไป (next label) ของตัวแปรร่วมเป็นจุดเทียบ

การเปรียบเทียบแบบใช้ระดับกลุ่มก่อนหน้า หรือ ถัดไปของตัวแปรร่วมเป็นจุดเทียบ ด้วยคำสั่งหลัก **margins** ภายใต้ตัวเลือกคำสั่ง **contrast** เป็นวิธีการพิจารณาการเปรียบเทียบค่าผลกระทบภายใต้แนวคิด ดังนี้

“ค่าผลกระทบที่ได้ในแต่ละระดับ ถูกลำมาเปรียบเทียบรายคู่กับค่าระดับกลุ่มก่อนหน้า (previous label) หรือ ถัดไป (next level) ของตัวแปรร่วมที่นำมาพิจารณา”

จากกรณีตัวอย่างที่ 1.2 โดยใช้ผลการวิเคราะห์ถดถอยลอจิสติกเช่นเดิม นักวิจัยสามารถเปรียบเทียบค่าผลกระทบแบบใช้ระดับกลุ่มก่อนหน้า (previous label) ของตัวแปรร่วมเป็นจุดเทียบ โดยการเพิ่ม “**ar.**” หน้าตัวแปรร่วมที่นำมาพิจารณา หรือ เปรียบเทียบค่าผลกระทบแบบใช้ระดับกลุ่มถัดไป (next label) ของตัวแปรร่วมเป็นจุดเทียบ โดยการเพิ่ม “**a.**” หน้าตัวแปรร่วมที่นำมาพิจารณา ซึ่งมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.19



แผนภาพที่ 1.19 แสดงผลลัพธ์ที่ได้กรณีเปรียบเทียบค่าผลกระทบแบบใช้ระดับกลุ่มก่อนหน้า หรือ ถัดไปเป็นจุดเทียบ

1.4.4 การเปรียบเทียบแบบรายคู่ที่เป็นไปได้ทั้งหมด

เป็นการเปรียบเทียบแบบรายคู่ที่เป็นไปได้ทั้งหมดของค่าประมาณผลกระทบ เช่นเดียวกับ pairwise comparison หรือ post hoc test ด้วยคำสั่งหลัก **margins** ภายใต้ตัวเลือกคำสั่ง **pwcompare** ซึ่งเป็นวิธีการพิจารณาการเปรียบเทียบค่าผลกระทบภายใต้แนวคิดดังนี้

“ค่าผลกระทบที่ได้ในแต่ละระดับ ถูกนำมาเปรียบเทียบรายคู่ ที่เป็นไปได้ทั้งหมด”

จากกรณีตัวอย่างที่ 1.2 โดยใช้ผลการวิเคราะห์ถดถอยลอจิสติกเช่นเดิม นักวิจัยสามารถเปรียบเทียบค่าผลกระทบแบบรายคู่ที่เป็นไปได้ทั้งหมดของตัวแปรพร้อมที่นำมาพิจารณา ซึ่งมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.20

```
. quietly logit died i.drug age
. margins drug,pwcompare
```

Pairwise comparisons of predictive margins
Model VCE : OIM

Expression : Pr(died), predict()

		Delta-method		Unadjusted	
		Contrast	Std. Err.	[95% Conf. Interval]	

	drug				
2	vs 1	-.5506517	.1355682	-.8163604	-.284943
3	vs 1	-.4841477	.1359077	-.750522	-.2177734
3	vs 2	.066504	.1801838	-.2866497	.4196577

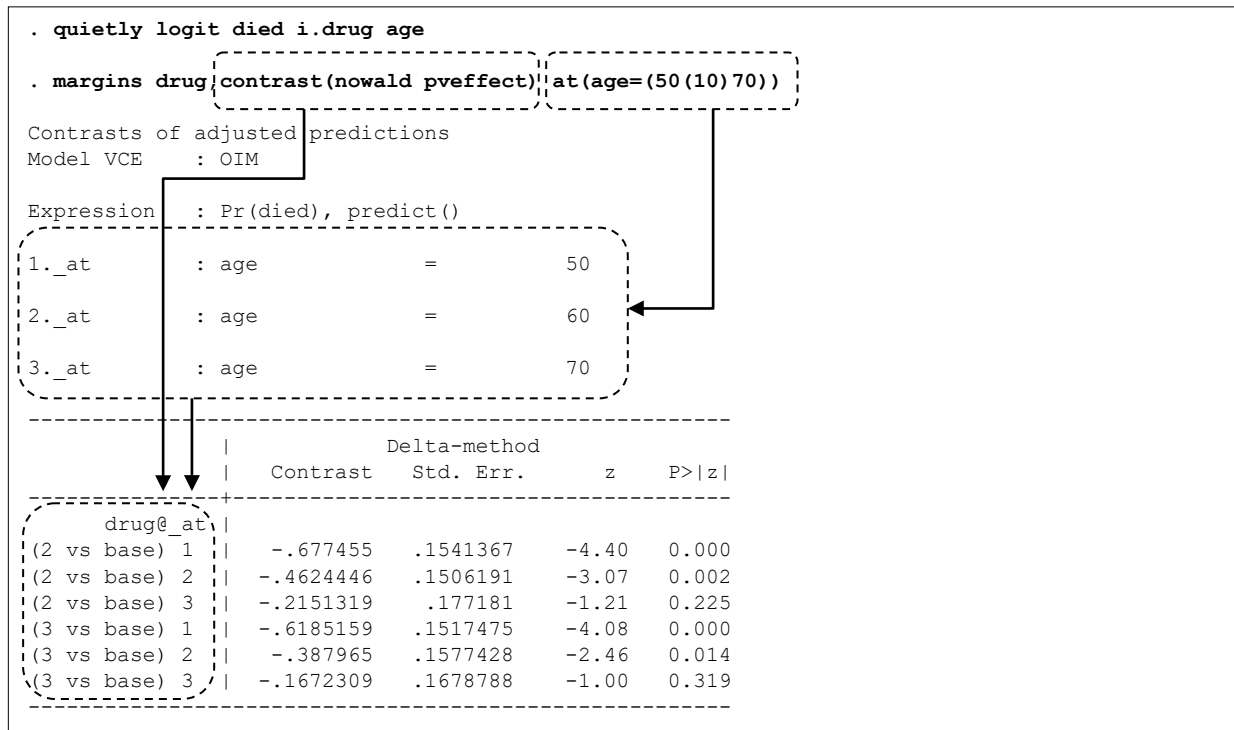
แผนภาพที่ 1.20 แสดงผลลัพธ์ที่ได้กรณีเปรียบเทียบค่าผลกระทบแบบรายคู่ที่เป็นไปได้ทั้งหมด

1.4.5 การเปรียบเทียบแบบใช้ระดับพื้นฐานของตัวแปรร่วมเป็นจุดเทียบ ณ ตำแหน่งที่นักวิจัยสนใจ

การเปรียบเทียบค่าประมาณผลกระทบแบบใช้ระดับพื้นฐานของตัวแปรร่วมเป็นจุดเทียบ ณ ตำแหน่งที่นักวิจัยสนใจ ด้วยคำสั่งหลัก **margins** ภายใต้ตัวเลือกคำสั่ง **contrast** เป็นวิธีการพิจารณาการเปรียบเทียบค่าผลกระทบภายใต้แนวคิดดังนี้

“ค่าผลกระทบที่ได้ในแต่ละระดับ ถูกนำมาเปรียบเทียบกับค่าระดับพื้นฐาน (base level) หรือค่าระดับแรกสุดของตัวแปรพร้อมที่นำมาพิจารณา พร้อมกับการกำหนดค่าตัวแปรอื่นประกอบด้วยการระบุตำแหน่งที่สนใจและต้องการแสดงผล”

จากกรณีตัวอย่างที่ 1.2 และใช้ผลการวิเคราะห์ถดถอยลอจิสติกเช่นเดิม นักวิจัยสามารถเปรียบเทียบค่าผลกระทบแบบใช้ระดับพื้นฐานของตัวแปรร่วมเป็นจุดเทียบ ณ ตำแหน่งที่นักวิจัยสนใจ ซึ่งมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.21



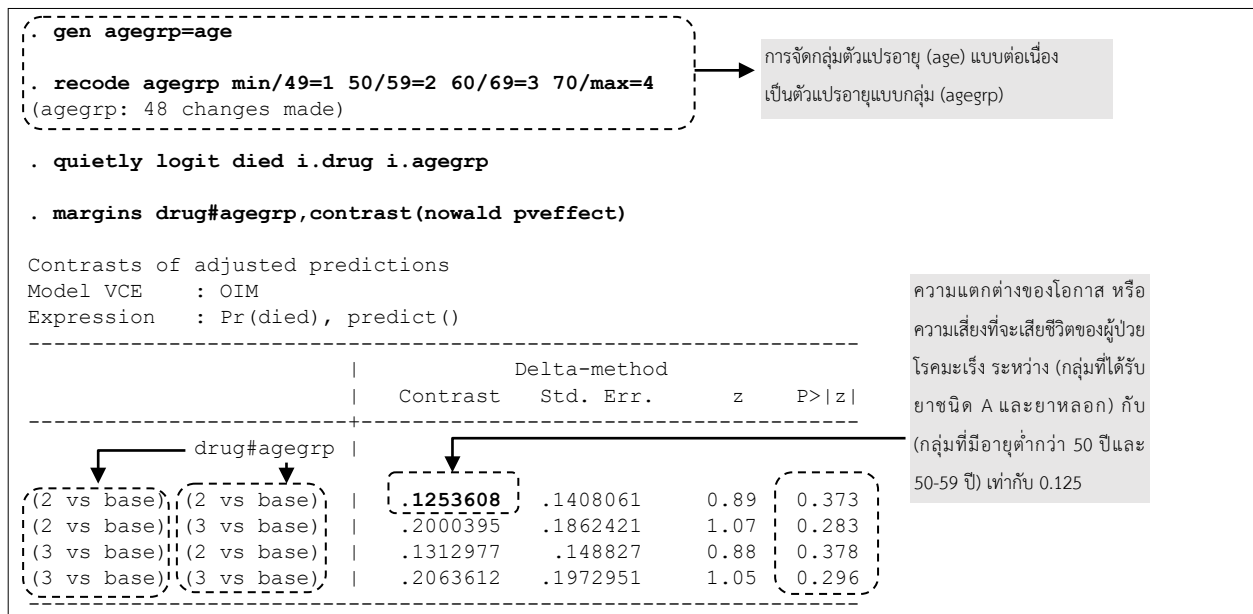
แผนภาพที่ 1.21 แสดงผลลัพธ์ที่ได้กรณีเปรียบเทียบโดยใช้ระดับพื้นฐานตัวแปรร่วมเป็นจุดเทียบ ณ ตำแหน่งที่สนใจ

1.4.6 การเปรียบเทียบแบบรายคู่ระหว่างสองตัวแปรร่วม

การเปรียบเทียบค่าประมาณผลกระทบแบบรายคู่ระหว่างสองตัวแปรร่วม ด้วยคำสั่งหลัก **margins** ภายใต้ตัวเลือกคำสั่ง **contrast** เป็นวิธีการพิจารณาการเปรียบเทียบค่าผลกระทบภายใต้แนวคิดดังนี้

“แม้ตัวแปรร่วมระหว่าง 2 ตัวแปร จะไม่มีปฏิสัมพันธ์ร่วม (no interaction) แต่ก็สามารถนำมาพิจารณาเปรียบเทียบผลต่างของผลกระทบในแต่ละระดับได้ ภายใต้หลักการเดิมของการเปรียบเทียบรายคู่กับค่าระดับพื้นฐาน (base level) หรือ ค่าระดับแรกสุดของตัวแปรร่วมทั้งสองที่นำมาพิจารณา”

จากกรณีตัวอย่างที่ 1.2 พิจารณาจัดกลุ่มตัวแปรอายุ (age) เป็นตัวแปรกลุ่ม (agegrp) ประกอบด้วย 1=ต่ำกว่า 50 ปี, 2=50-59 ปี, 3=60-69 ปี และ 4=70 ปีขึ้นไป โดยใช้คำสั่ง **gen** และคำสั่ง **recode** และใช้วิธีการวิเคราะห์ถดถอยลอจิสติกเช่นเดิม นักวิจัยสามารถเปรียบเทียบค่าผลกระทบแบบรายคู่ระหว่างสองตัวแปรร่วม โดยมีคำสั่งและผลลัพธ์ที่ได้ ดังแผนภาพที่ 1.22



แผนภาพที่ 1.22 แสดงผลลัพธ์ที่ได้กรณีเปรียบเทียบโดยใช้ระดับพื้นฐานตัวแปรร่วมเป็นจุดเทียบ ณ ตำแหน่งที่สนใจ

จากแผนภาพที่ 1.22 พบว่า เป็นการเปรียบเทียบค่าผลกระทบที่เกิดขึ้นระหว่างสองตัวแปรร่วม ได้แก่ ตัวแปร drug (1=ยาหลอก, 2=ยาชนิด A และ 3=ยาชนิด B) และตัวแปร agegrp (1=ต่ำกว่า 50 ปี, 2=50-59 ปี, 3=60-69 ปีและ 4=70 ปีขึ้นไป) แม้ว่า ตัวแปรทั้งสองมีปฏิกริยาร่วมกัน (interaction) หรือไม่ ? ก็ตาม โดยจะพบว่า ในแต่ละรายคู่จะมีการเปรียบเทียบ 2 ครั้ง นั่นคือ ครั้งที่ 1 เป็นการเปรียบเทียบภายในของแต่ละตัวแปรทั้งสอง โดยมีจุดเทียบเป็นระดับพื้นฐาน (base level) ของตัวแปรเหล่านั้น ๆ ครั้งที่ 2 เป็นการนำเอาค่าผลกระทบที่ได้จากการเปรียบเทียบภายในตัวแปรร่วม มาเปรียบเทียบอีกครั้งระหว่างตัวแปรทั้งสอง

1.5 บทสรุป

ในทางปฏิบัติ คำสั่ง **margins** จะถูกนำมาใช้เมื่อนักวิจัยต้องการประมาณ หรือ ทำนายค่าผลกระทบที่เกิดขึ้น โดยผลลัพธ์ที่ได้ จะประกอบด้วย ค่าประมาณ ซึ่งเป็นผลลัพธ์จากการทำนายค่า (prediction) ภายใต้สมการตัวแบบที่ถูกนำมาวิเคราะห์ข้อมูล ขณะที่คำสั่ง **contrast** จะถูกนำมาใช้เมื่อนักวิจัยต้องการเปรียบเทียบความแตกต่างของผลกระทบที่เกิดขึ้นแบบรายคู่ โดยผลลัพธ์ที่ได้ จะประกอบด้วย ผลต่างที่เกิดขึ้นจากการเปรียบเทียบ นอกจากนี้คำสั่งทั้งสองดังกล่าว ยังมีตัวเลือกอื่น ที่สามารถนำมากำหนด เพื่อแสดงผลลัพธ์ตามที่ต้องการได้ ซึ่งในรายละเอียดดังกล่าว นักวิจัยสามารถเรียนรู้และศึกษาเพิ่มเติมได้จากระบบความช่วยเหลือของโปรแกรม STATA ด้วยการพิมพ์คำสั่ง **help margins** หรือ **help contrast**