

ผลการวิเคราะห์ข้อมูล

3.1 การหาตัวแทนของครัวเรือนที่ไม่ตอบในเรื่องรายได้

เนื่องจากหน่วยตัวอย่างที่ไม่ตอบในเรื่องรายได้ ซึ่งนำมาจากโครงการวิจัยเรื่องสภาวะทางเศรษฐกิจ สังคม และประชากรของกรุงเทพมหานคร ในทางปฏิบัติเราไม่สามารถหาคำรายได้ที่แท้จริงของหน่วยตัวอย่างดังกล่าวจากแหล่งอื่นได้ด้วย เหตุนี้จึงไม่สามารถนำตัวอย่างเหล่านี้มาวิเคราะห์ รูปแบบความสัมพันธ์ของความน่าจะเป็นในการตอบเรื่องรายได้ในกรณีที่ความน่าจะเป็นของการตอบรายได้ขึ้นอยู่กับระดับรายได้ของผู้ตอบ กล่าวคือ ลักษณะตัวอย่างที่นำมาศึกษาครั้งนี้ เราไม่สามารถทดสอบได้ว่ารายได้ที่ไม่ได้รับคำตอบขึ้นอยู่กับระดับรายได้หรือไม่ ดังนั้นในการวิจัยครั้งนี้จึงหาตัวแทนของครัวเรือนที่ไม่ตอบในเรื่องรายได้ จากครัวเรือนที่ให้คำตอบเรื่องรายได้ เพื่อหาวิธีที่เหมาะสมสำหรับการประมาณค่ารายได้ที่ขาดหายไป

โดยอาศัยข้อมูลจากโครงการวิจัย เรื่อง สภาวะทางเศรษฐกิจ สังคม และประชากรของกรุงเทพมหานคร ดังกล่าวมาแล้วข้างต้น จากตัวอย่างทั้งหมด 3,306 ครัวเรือน เมื่อคัดหน่วยตัวอย่างที่ให้ค่าตัวแปรทุกตัวจากตัวแปรที่นำมาใช้ในการวิจัยครั้งนี้ ยกเว้นตัวแปรรายได้เพียงตัวเดียวที่อาจได้รับคำตอบหรือไม่ได้รับคำตอบก็ได้ และคัดเลือกเฉพาะหัวหน้าครัวเรือนที่ได้รับคำตอบแทนจากอาชีพหลักเท่านั้น ดังได้อธิบายมาแล้วในบทที่ 1 จะเหลือจำนวนตัวอย่าง 1,513 ครัวเรือน ในจำนวนนี้มีพวกที่ไม่ตอบเรื่องรายได้ 77 ครัวเรือน ดังนั้นจะเหลือครัวเรือนที่ให้คำตอบเรื่องรายได้ 1,436 ครัวเรือน สำหรับใช้ในการหาตัวแทนครัวเรือนที่ไม่ตอบเรื่องรายได้ เพื่อศึกษาวิธีการหาค่ารายได้ที่ขาดหายไปเป็น 2 กรณีคือ

1. เลือกตัวแทนโดยอาศัยปัจจัยที่คิดว่ามีอิทธิพลต่อรายได้หลักของหัวหน้าครัวเรือนตามที่ Greenlees and others (1982) ได้ศึกษาแล้วว่า ตัวแปรที่มีอิทธิพลต่อการตอบเรื่องรายได้ คือ รายได้ อายุ การสัมภาษณ์โดยส่งพนักงานไปสัมภาษณ์ การศึกษา ผิด ภาคที่ครัวเรือนตั้งอยู่ สำหรับการเลือกตัวแทนด้วยวิธีนี้ เป็นการสมมติว่าไม่ทราบรายได้ของตัวแทนชุดนี้ เพื่อศึกษาว่าถ้าตัวแทนชุดนี้ไม่ให้คำตอบเรื่องรายได้ เนื่องจากสาเหตุต่าง ๆ ตามเงื่อนไขที่สร้างขึ้นแล้วจะมีผลต่อการใช้วิธีการประมาณค่ารายได้ที่ขาดหายไป ซึ่งจะนำมาใช้ในการวิจัยนี้อย่างไร

2. เลือกตัวแทนโดยสุ่ม และสมมติว่า ไม่ทราบรายได้ของตัวแทนชุดนี้ เพื่อศึกษาดูว่า ถ้าการไม่  
ให้คำตอบ เรื่องรายได้ของตัวแทนชุดดังกล่าว เป็นไปโดยบังเอิญมากกว่าตั้งใจ จะมีผลต่อการใช้วิธีการ  
ประมาณค่ารายได้ที่ขาดหายไปซึ่งจะนำมาใช้ในการวิจัยนี้อย่างไร

สำหรับการเลือกตัวแทนทั้ง 2 กรณี เนื่องจาก Greenlees and others (1982: 252)  
ได้กล่าวว่า ถ้ารายได้ที่ไม่ได้รับคำตอบ เนื่องจากความน่าจะเป็นของการตอบขึ้นอยู่กับระดับรายได้แล้ว  
การหาค่ารายได้ที่ขาดหายไปด้วยวิธีสมการถดถอยจะให้ผลการประมาณค่ารายได้ที่ไม่น่าเชื่อถือ ด้วย  
เหตุนี้จึงศึกษาดูว่าวิธีประมาณค่ารายได้ที่ขาดหายไปโดยการแบ่งชั้นภูมิ เมื่อเลือกตัวอย่างแล้วจะใช้ได้  
ทั้ง 2 กรณีข้างต้นหรือไม่ และถ้าวิธีประมาณค่าด้วยวิธีสมการถดถอยไม่ได้ผลในกรณีที่ตัวอย่างไม่ได้  
ขาดหายไปโดยสุ่มจะเป็นไปตามที่ Greenlees and others สรุปไว้หรือไม่

### 3.2 ผลการหาตัวแทนของครัวเรือนที่ไม่ตอบในเรื่องรายได้

ในการหาตัวแทนของครัวเรือนที่ไม่ตอบในเรื่องรายได้จำนวน 77 ครัวเรือนได้ผลดังนี้

1) กรณีที่เลือกตัวแทนโดยอาศัยปัจจัยพื้นฐานที่คิดว่ามีอิทธิพลต่อรายได้หลักของหัวหน้า  
ครัวเรือน ปัจจัยที่นำมาพิจารณาในที่นี้ได้ใช้ อายุ การศึกษา ตามที่ Greenlees and others  
ได้ศึกษาไว้ นอกจากนี้ยังใช้ปัจจัยอื่น ๆ เพิ่มเติมโดยพิจารณาจากความสัมพันธ์ระหว่างระดับรายได้  
หลักของหัวหน้าครัวเรือนกับปัจจัยอื่น ๆ ซึ่งได้แก่ อาชีพหลัก สถานภาพในการทำงานอาชีพหลัก  
ฐานะทางเศรษฐกิจของผู้ให้สัมภาษณ์ รายจ่ายสำหรับครัวเรือน จำนวนชั่วโมงทำงานสำหรับอาชีพ  
หลัก ประสบการณ์ในการทำงานอาชีพหลัก เขตที่ครัวเรือนตั้งอยู่

ขั้นตอนในการเลือกตัวแทนด้วยวิธีข้างต้น มีดังต่อไปนี้

(1) พิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับปัจจัยอื่น ๆ  
ดังกล่าวมาแล้ว โดยใช้การทดสอบไคสแควร์โดยตั้งสมมติฐานว่าง ว่าไม่มีความสัมพันธ์ระหว่างตัวแปร

2. ตัว สูตรที่ใช้ในการคำนวณเป็นดังนี้

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$$

เมื่อ k แทนจำนวนเซลล์ทั้งหมด

$O_i$  แทน ค่าความถี่ที่สังเกตได้ (Observed frequency)

ในแต่ละเซลล์

- $E_i$  แทน ค่าความถี่ที่คาดหวัง (expected frequency)
- $r$  และ  $c$  แทน จำนวนแถวและสดมภ์ตามลำดับ
- $N$  แทน ผลรวมของความถี่ทั้งหมด
- $\chi^2 = (r-1)(c-1)$  แทน จำนวนองศาแห่งความเป็นอิสระ
- ถ้า  $\chi^2 > \chi_{\alpha}^2$  จะปฏิเสธสมมติฐานว่างที่ว่า ตัวแปรทั้งสองเป็นอิสระกันด้วยระดับ

นัยสำคัญ  $\alpha$

ผลการทดสอบไคสแควร์ ของตัวแปรรายได้กับตัวแปรอื่น ๆ คือ อายุ การศึกษา อาชีพหลัก สถานภาพในการทำงานสำหรับอาชีพหลัก จำนวนชั่วโมงทำงานต่อสัปดาห์สำหรับอาชีพหลัก ประสบการณ์ในการทำงาน อาชีพหลัก ของหัวหน้าครัวเรือน ฐานะทางเศรษฐกิจของผู้ให้สัมภาษณ์ รายจ่ายสำหรับครัวเรือน และเขตที่ครัวเรือนตั้งอยู่ ได้แสดงไว้ในตารางที่ 1 ถึงตารางที่ 9

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 1 จำนวนครัวเรือนจำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและระดับอายุของหัวหน้าครัวเรือน

รายได้หลัก (บาท/เดือน)	อายุ (ปี)							รวม
	ต่ำกว่า 30	30-34	35-39	40-44	45-49	50-54	มากกว่า 55	
ต่ำกว่า 2,000	77	73	60	56	44	50	54	414
2,001-4,000	101	99	83	75	73	43	37	511
4,001-6,000	25	41	34	43	24	25	25	217
6,001-8,000	8	19	23	15	13	11	7	96
8,001-10,000	12	11	13	13	15	8	15	87
10,001-15,000	5	6	9	4	9	7	13	53
มากกว่า 15,000	5	8	9	15	5	8	8	58
รวม	233	257	231	221	183	152	159	1,436

$$\chi^2 = 66.514$$

$$d.f. = 36$$

จากตารางที่ 1 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับระดับอายุของหัวหน้าครัวเรือน การทดสอบสมมติฐานว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับระดับอายุของหัวหน้าครัวเรือน ผลจากการทดสอบไคสแควร์ได้ค่า  $\chi^2$  เท่ากับ 66.514 ด้วยระดับของเสรีภาพ 36 ดังนั้นที่ระดับนัยสำคัญ .05 จะต้องปฏิเสธสมมติฐานที่ว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับระดับอายุของหัวหน้าครัวเรือน กล่าวคือหัวหน้าครัวเรือนที่มีระดับอายุต่างกัน มีลักษณะการแจกแจงของรายได้ต่างกัน



ตารางที่ 2 จำนวนครัวเรือน จำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและระดับการศึกษา

รายได้หลัก (บาท/เดือน)	ระดับการศึกษา				
	ไม่เคยศึกษา	ประถมศึกษา	มัธยมศึกษา	ตั้งแต่ปริญญาตรี ขึ้นไป	รวม
ต่ำกว่า 2,000	42	311	54	7	41
2,001-4,000	33	298	129	51	511
4,001-6,000	14	101	55	47	217
6,001-8,000	6	35	36	19	96
8,001-10,000	5	37	24	21	87
10,001-15,000	7	18	7	21	53
มากกว่า 15,000	2	27	21	8	58
รวม	109	827	326	174	1,436

$$\chi^2 = 192.611$$

$$d.f. = 18$$

จากตารางที่ 2 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือน กับระดับการศึกษา การทดสอบสมมติฐานว่า ระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับระดับการศึกษา ผลจากการทดสอบไคสแควร์ได้ค่า  $\chi^2$  เท่ากับ 192.611 ด้วยระดับองศาอิสระ 18 ดังนั้นที่ระดับนัยสำคัญ .05 จะต้องปฏิเสธสมมติฐานที่ว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับระดับการศึกษา กล่าวคือ หัวหน้าครัวเรือนที่มีระดับการศึกษาต่างกัน จะมีลักษณะการแจกแจงของรายได้ต่างกัน

ตารางที่ 3 จำนวนครัวเรือนจำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและอาชีพหลัก

รายได้หลัก (บาท/เดือน)	อาชีพหลัก								รวม
	1	2	3	4	5	6	7	8	
ต่ำกว่า 2,000	8	2	13	138	52	5	155	41	414
2,001-4,000	33	6	71	149	20	25	164	43	511
4,001-6,000	40	5	28	79	5	10	41	9	217
6,001-8,000	18	2	16	35	2	6	15	2	96
8,001-10,000	16	5	6	38	2	1	14	5	87
10,001-15,000	15	7	1	20	3	2	3	3	54
มากกว่า 15,000	4	7	2	35	4	1	3	1	57
รวม	134	34	137	494	88	50	395	104	1,436

$$\chi^2 = 332.762$$

$$d.f. = 42$$

จากตารางที่ 3 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับอาชีพหลัก การทดสอบสมมติฐานว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับอาชีพหลักของหัวหน้าครัวเรือน ผลจากการทดสอบไคสแควร์ได้ค่า  $\chi^2$  เท่ากับ 332.762 ด้วยระดับองศาอิสระ 42 ดังนั้นที่ระดับนัยสำคัญ .05 จะต้องปฏิเสธสมมติฐานที่ว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับอาชีพหลัก กล่าวคือ หัวหน้าครัวเรือนที่มีลักษณะของอาชีพหลักต่างกัน จะมีลักษณะการแจกแจงของรายได้ต่างกัน

หมายเหตุ .

- |   |  |
|---|--|
| 1 หมายถึง วิชาชีพ เทคนิค และที่เกี่ยวข้อง | 5 หมายถึง เกษตร ประมง ล่าสัตว์และที่เกี่ยวข้อง |
| 2 หมายถึง อำนวยการ บริหารและการจัดการ     | 6 หมายถึง การคมนาคมขนส่ง                       |
| 3 หมายถึง เสมียนพนักงาน                   | 7 หมายถึง แกะสลักและการผลิตอื่น                |
| 4 หมายถึง การขาย                          | 8 หมายถึง บริการ กีฬา และสันทนาการ             |

ตารางที่ 4 จำนวนครัวเรือนจำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและสถานภาพในการทำงานสำหรับอาชีพหลัก

รายได้หลัก (บาท/เดือน)	สถานภาพในการทำงานอาชีพหลัก						รวม
	1	2	3	4	5	6	
ต่ำกว่า 2,000	109	4	113	29	4	155	414
2,001-4,000	123	13	120	103	42	110	511
4,001-6,000	71	6	35	46	21	38	217
6,001-8,000	32	2	19	17	17	9	96
8,001-10,000	45	2	14	11	4	11	87
10,001-15,000	25	1	7	10	2	7	52
มากกว่า 15,000	38	4	6	1	2	8	59
รวม	443	32	314	217	92	338	1,436

$$\chi^2 = 220.184$$

$$d.f. = 30$$

จากตารางที่ 4 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือน กับสถานภาพในการทำงานสำหรับอาชีพหลัก การทดสอบสมมติฐานว่า ระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับสถานภาพในการทำงานสำหรับอาชีพหลัก ผลจากการทดสอบไคสแควร์ ได้ค่า  $\chi^2$  เท่ากับ 220.184 ด้วยระดับของคาอิสระ 30 ดังนั้นที่ระดับนัยสำคัญ .05 จะต้องปฏิเสธสมมติฐานที่ว่า ระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับสถานภาพในการทำงานสำหรับอาชีพหลัก กล่าวคือหัวหน้าครัวเรือนที่มีสถานภาพในการทำงานสำหรับอาชีพหลักต่างกันจะมีการแจกแจงของรายได้ต่างกัน

#### หมายเหตุ

- |                                 |  |
|---------------------------------|--|
| 1 หมายถึง เป็นเจ้าของกิจการ     | 4 หมายถึง เป็นข้าราชการ/ลูกจ้างรัฐบาล    |
| 2 หมายถึง เป็นหุ้นส่วนของกิจการ | 5 หมายถึง เป็นลูกจ้าง/พนักงานรัฐวิสาหกิจ |
| 3 หมายถึง เป็นลูกจ้างเอกชน      | 6 หมายถึง ทำงานส่วนตัว                   |

ตารางที่ 5 จำนวนครัวเรือนจำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและฐานะทางเศรษฐกิจของหัวหน้าครัวเรือน

รายได้หลัก (บาท/เดือน)	ฐานะทางเศรษฐกิจของหัวหน้าครัวเรือน			
	ดี	ปานกลาง	ไม่ดี	รวม
ต่ำกว่า 2,000	29	264	121	414
2,001-4,000	62	370	79	511
4,001-6,000	70	133	14	217
6,001-8,000	33	61	2	96
8,001-10,000	47	39	1	87
10,001-15,000	33	17	3	53
มากกว่า 15,000	40	18	0	58
รวม	314	902	220	1,436

$$\chi^2 = 343.441$$

$$d.f. = 12$$

จากตารางที่ 5 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับฐานะทางเศรษฐกิจของหัวหน้าครัวเรือน (เป็นการประเมินอย่างกว้าง ๆ โดยพนักงานสัมภาษณ์ ซึ่งพิจารณาจากสภาพของบ้าน สิ่งอำนวยความสะดวกภายในบ้าน และอื่น ๆ โดยรวม แต่ไม่ได้กำหนดหลักเกณฑ์เอาไว้ตายตัว) การทดสอบสมมติฐานที่ว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับฐานะทางเศรษฐกิจของหัวหน้าครัวเรือน ผลจากการทดสอบไคสแควร์ได้ค่า  $\chi^2 = 343.441$  ด้วยระดับของเสรีอิสระ 12 ดังนั้นที่ระดับนัยสำคัญ .05 จะต้องปฏิเสธสมมติฐานที่ว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับฐานะทางเศรษฐกิจของหัวหน้าครัวเรือน กล่าวคือ หัวหน้าครัวเรือนที่มีฐานะทางเศรษฐกิจต่างกัน มีลักษณะการแจกแจงของรายได้ต่างกัน



ตารางที่ 6 จำนวนครัวเรือนจำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและระดับรายจ่ายสำหรับครัวเรือน

รายได้หลัก (บาท/เดือน)	รายจ่ายสำหรับครัวเรือน (บาท/เดือน)							รวม
	ต่ำกว่า 2,000	2,001- 4,000	4,001- 6,000	6,001- 8,000	8,001- 10,000	10,001- 15,000	มากกว่า 15,000	
ต่ำกว่า 2,000	143	196	42	12	8	8	5	414
2,001-4,000	88	277	98	26	15	4	3	511
4,001-6,000	20	78	68	14	21	10	6	217
6,001-8,000	2	23	33	16	12	4	6	96
8,001-10,000	3	13	30	17	18	3	3	87
10,001-15,000	5	5	8	9	15	7	4	53
มากกว่า 15,000	2	4	8	5	17	8	14	58
รวม	263	596	287	99	106	44	41	1,436

$$\chi^2 = 579.303$$

$$d.f. = 36$$

จากตารางที่ 6 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับระดับรายจ่ายสำหรับครัวเรือน การทดสอบสมมติฐานว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับระดับรายจ่ายสำหรับครัวเรือน ผลจากการทดสอบไคสแควร์ได้ค่า  $\chi^2 = 579.303$  ด้วยระดับองศาอิสระ 36 ดังนั้นที่ระดับนัยสำคัญ .05 จะต้องปฏิเสธสมมติฐานที่ว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับระดับรายจ่ายสำหรับครัวเรือน กล่าวคือหัวหน้าครัวเรือนที่มีระดับรายจ่ายต่างกันจะมีลักษณะการแจกแจงของรายได้ต่างกัน

ตารางที่ 7 จำนวนครัวเรือน จำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและจำนวน ชั่วโมงทำงานต่อสัปดาห์สำหรับอาชีพหลัก

รายได้หลัก (บาท/เดือน)	จำนวนชั่วโมงทำงานต่อสัปดาห์สำหรับอาชีพหลัก				
	ต่ำกว่าหรือ เท่ากับ 30	31-40	41-50	มากกว่า 50	รวม
ต่ำกว่า 2,000	69	67	97	181	414
2,001-4,000	45	125	127	214	511
4,001-6,000	29	56	43	89	217
6,001-8,000	6	31	20	39	96
8,001-10,000	7	20	12	48	87
10,001-15,000	5	18	8	22	53
มากกว่า 15,000	7	8	6	37	58
รวม	168	325	313	630	1,436

$$\chi^2 = 54.083$$

$$d.f. = 18$$

จากตารางที่ 7 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือน กับจำนวนชั่วโมงทำงานต่อสัปดาห์สำหรับอาชีพหลักสำหรับการทดสอบสมมติฐานว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับจำนวนชั่วโมงทำงานต่อสัปดาห์สำหรับอาชีพหลัก ผลจากการทดสอบไคสแควร์ได้ค่า  $\chi^2$  เท่ากับ 54.083 ด้วยระดับองศาอิสระ 18 จะต้องปฏิเสธสมมติฐานที่ว่าระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับจำนวนชั่วโมงทำงานต่อสัปดาห์ กล่าวคือ หัวหน้าครัวเรือนที่มีจำนวนชั่วโมงทำงานต่อสัปดาห์สำหรับอาชีพหลักต่างกันจะมีลักษณะการแจกแจงของรายได้ต่างกัน

**ตารางที่ 8** จำนวนครัวเรือนจำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและ  
ประสบการณ์ในการทำงานสำหรับอาชีพหลัก

รายได้หลัก (บาท/เดือน)	ประสบการณ์ในการทำงานสำหรับอาชีพหลัก (ปี)						
	ต่ำกว่า 10	10-14	15-20	21-30	31-40	มากกว่า 40	รวม
ต่ำกว่า 2,000	28	21	68	117	87	93	414
2,001-4,000	70	37	72	144	116	72	511
4,001-6,000	24	24	34	50	49	36	217
6,001-8,000	8	10	17	31	21	9	96
8,001-10,000	6	11	11	26	15	18	87
10,001-15,000	5	3	9	12	16	8	53
มากกว่า 15,000	2	5	13	12	17	9	58
รวม	143	111	224	392	321	245	1,436

$$\chi^2 = 51.965$$

$$d.f. = 30$$

จากตารางที่ 8 แสดงจำนวนครัวเรือนจำแนกตามระดับรายได้หลักและประสบการณ์  
ในการทำงานสำหรับอาชีพหลักของหัวหน้าครัวเรือน สำหรับการทดสอบสมมติฐานว่าระดับรายได้  
หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับประสบการณ์ในการทำงานสำหรับอาชีพหลัก ผลจาก  
การทดสอบไคสแควร์ได้ค่า  $\chi^2 = 51.965$  ด้วยระดับของเสรีอิสระ 30 จะต้องปฏิเสธสมมติฐานที่ว่า  
ระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับประสบการณ์ในการทำงานสำหรับอาชีพหลัก  
กล่าวคือ หัวหน้าครัวเรือนที่มีประสบการณ์ในการทำงานต่างกัน จะมีลักษณะการแจกแจงของรายได้  
ต่างกัน ที่ระดับนัยสำคัญ .05

ตารางที่ 9 จำนวนครัวเรือนจำแนกตามระดับรายได้หลักของหัวหน้าครัวเรือนและ เขตที่ครัวเรือนอาศัยอยู่

รายได้หลัก (บาท/เดือน)	เขต			
	เขตเมือง	เขตต่อเมือง	เขตชานเมือง	รวม
ต่ำกว่า 2,000	180	163	71	414
2,001-4,000	245	204	62	511
4,001-6,000	113	86	18	217
6,001-8,000	41	48	7	96
8,001-10,000	41	38	8	87
10,001-15,000	18	31	4	53
มากกว่า 15,000	37	15	6	58
รวม	675	585	176	1,436

$$\chi^2 = 32.54671$$

$$d.f. = 12$$

จากตารางที่ 9 ถ้าพิจารณาความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับเขตที่ครัวเรือนอาศัยอยู่ การทดสอบสมมติฐานว่า ระดับรายได้หลักของหัวหน้าครัวเรือนไม่มีความสัมพันธ์กับเขตที่ครัวเรือนอาศัยอยู่ ผลจากการทดสอบไคสแควร์ได้ค่า  $\chi^2 = 32.54671$  ด้วยระดับองศาอิสระ 12 สรุปได้ว่า ระดับรายได้หลักของหัวหน้าครัวเรือนมีความสัมพันธ์กับเขตที่ครัวเรือนอาศัยอยู่ กล่าวคือ หัวหน้าครัวเรือนที่อาศัยอยู่ในเขตต่าง ๆ กันจะมีลักษณะการแจกแจงของรายได้ต่างกันที่ระดับนัยสำคัญ .05

เมื่อพิจารณาถึง ความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับปัจจัยอื่น ๆ ข้างต้น ผลจากตารางที่ 1 ถึงตารางที่ 9 สรุปได้ว่าปัจจัยต่าง ๆ มีความสัมพันธ์กับระดับรายได้หลักของหัวหน้าครัวเรือนทุกปัจจัย

(2) เลือกปัจจัยพื้นฐานสำหรับหัวหน้าครัวเรือน เพื่อใช้เป็นหลักในการเลือกตัวแทนให้มีลักษณะคล้ายกับตัวอย่างที่ไม่ตอบทั้ง 77 ครัวเรือน กล่าวคือมีการตั้งเงื่อนไขโดยใช้ลักษณะของครัวเรือนที่ไม่ตอบทั้ง 77 ครัวเรือน มาหาตัวแทนจากครัวเรือนที่เหลืออยู่ 1,436 ครัวเรือนสำหรับครัวเรือนที่เป็นตัวแทนนี้ จะมีค่าตอบเกี่ยวกับรายได้ทุกครัวเรือนโดยมีการสมมติว่าตัวแทนของครัวเรือน 77 ครัวเรือน เป็นครัวเรือนที่ไม่ให้คำตอบเรื่องรายได้ เนื่องจากอิทธิพลของรายได้และปัจจัยอื่น ๆ ตามเงื่อนไขที่ตั้งไว้ใน การเลือกตัวแทน ปัจจัยที่ใช้หาลักษณะของหน่วยที่ตอบเพื่อแทนหน่วยที่ไม่ตอบเรื่องรายได้คือ อายุ การศึกษา อาชีพหลัก สถานภาพในการทำงานอาชีพหลักของหัวหน้าครัวเรือน ฐานะทางเศรษฐกิจ เขตที่ครัวเรือนอาศัยอยู่ จำนวนชั่วโมงทำงานต่อสัปดาห์ สำหรับอาชีพหลัก

(3) สัดเลือกตัวแทนโดยใช้ปัจจัยที่ได้จากข้อ (2) ที่แสดงลักษณะครัวเรือน ซึ่งไม่ตอบเรื่องรายได้ทั้ง 77 ครัวเรือน ดังแสดงไว้ในตาราง ก. 1 ของภาคผนวก ก. ตัวอย่างเช่น ลำดับที่ 1 ในตาราง ก. 1 แสดงหัวหน้าครัวเรือนที่มีอายุ 27 ปี การศึกษาจบชั้นมัธยมศึกษาปีที่ 3 ประกอบอาชีพแกะสลักและการผลิตอื่นและมีสถานภาพเป็นหุ้นส่วนของกิจการ ในเวลา 1 สัปดาห์ทำงาน 56 ชั่วโมง มีรายจ่ายสำหรับครัวเรือน 8,000 บาท ฐานะทางเศรษฐกิจที่พนักงานประเมินได้อยู่ในชั้นปานกลาง ครัวเรือนตั้งอยู่ในเขตเมือง จากลักษณะครัวเรือนที่ 1 นี้ ได้หาตัวแทนจากครัวเรือน 1,436 ครัวเรือน ที่ให้คำตอบเรื่องรายได้ โดยที่ครัวเรือนที่เลือกมามีลักษณะเหมือนครัวเรือนที่ 1 ในตาราง ก.1 และทำเช่นนี้กับครัวเรือนลำดับที่ 2 ไปจนถึงลำดับที่ 77 ในตาราง ก. 1 ลักษณะเดียวกับ ครัวเรือนที่ 1 สำหรับวิธีการคัดเลือกได้ใช้โปรแกรมสำเร็จรูป SPSS

(4) เนื่องจากจำนวนหน่วยที่ตอบและหน่วยที่ไม่ตอบในเรื่องรายได้มีจำนวนไม่ต่างกันมาก ทำให้โอกาสที่จะได้ครัวเรือนที่มีลักษณะเหมือนกับครัวเรือนที่ไม่ตอบในเรื่องรายได้เป็นไปได้น้อย ดังนั้นเมื่อทดลองหาตัวแทนตามวิธีในข้อ (3) จึงพบว่า ไม่อาจหาหน่วยที่ให้คำตอบเรื่องรายได้ จาก 1,436 ครัวเรือน เพื่อแทนครัวเรือนที่ไม่ตอบในเรื่องรายได้ทั้ง 77 ครัวเรือน

เพื่อให้ได้ตัวแทนที่จะแทนครัวเรือนทั้ง 77 ครัวเรือนให้มากที่สุด จึงได้กำหนดช่วงอย่างกว้าง ๆ ของปัจจัยบางปัจจัยและได้ตัดปัจจัยบางปัจจัยออก สำหรับการกำหนดช่วงตัวแปรบางตัวได้กำหนดตามความนิยมของผู้วิจัยโดยทั่วไปดังนี้

อายุ แบ่งกลุ่มช่วงอายุออกเป็น 10 กลุ่มคือ 15-19 ปี 20-24 ปี 25-29 ปี 30-34 ปี 35-39 ปี 40-44 ปี 45-49 ปี 50-54 ปี 55-64 ปี ตั้งแต่ 65 ปีขึ้นไป

การศึกษาแบ่งออกเป็น 5 ระดับคือ ไม่เคยศึกษา ประถมศึกษา มัธยมศึกษา ปริญญาตรี สูงกว่าปริญญาตรี

รายจ่ายสำหรับครัวเรือน แบ่งออกเป็น 12 กลุ่มคือ 10-2,000 บาท 2,001-4,000 บาท 4,001-6,000 บาท 6,001-8,000 บาท 8,001-10,000 บาท 10,001-15,000 บาท 15,001-20,000 บาท 20,001-25,000 บาท 25,001-30,000 บาท 30,001-40,000 บาท 40,001-50,000 บาท 50,001-100,000 บาท

ปัจจัยที่แสดงลักษณะของหน่วยที่ไม่ตอบทั้ง 77 ครัวเรือน ที่เหมาะสมคือ ระดับอายุ ระดับการศึกษา อาชีพหลัก รายจ่ายสำหรับครัวเรือน สถานภาพในการทำงานอาชีพหลัก ซึ่งสามารถหาหน่วยตัวอย่างได้ 61 ครัวเรือนแทนครัวเรือนที่ไม่ตอบเรื่องรายได้ 77 ครัวเรือน ดังได้แสดงไว้ตาราง ก. 2 ของภาคผนวก ก. และตัวอย่างที่ใช้ในการวิเคราะห์เพื่อหาค่าประมาณรายได้จากแต่ละวิธีจะเหลือ 1,375 ครัวเรือน จากจำนวนที่ตอบเรื่องรายได้ 1,436 ครัวเรือน

2) กรณีที่เลือกตัวแทนโดยวิธีสุ่ม จากหน่วยตัวอย่างที่ไม่ตอบเรื่องรายได้ 77 ครัวเรือนนั้น เมื่อคิดเทียบกับหน่วยตัวอย่างทั้งหมด 1,513 ครัวเรือน จะมีสัดส่วนของการไม่ตอบเท่ากับ 0.051 ดังนั้นได้ใช้โปรแกรมสำเร็จรูป SPSS หาตัวอย่างสุ่มที่มีสัดส่วน 0.051 จากครัวเรือนที่ให้คำตอบเรื่องรายได้ 1,436 ครัวเรือน จะได้ครัวเรือนตัวอย่างจำนวนทั้งสิ้น 74 ครัวเรือนเพื่อแทนครัวเรือน 77 ครัวเรือนที่ไม่ให้คำตอบเรื่องรายได้ และจะเหลือครัวเรือน 1,362 ครัวเรือน จากครัวเรือนที่ตอบเรื่องรายได้ 1,436 ครัวเรือน เพื่อใช้เป็นจำนวนตัวอย่างในการประมาณค่ารายได้ รายละเอียดของครัวเรือน 74 ครัวเรือนที่ให้คำตอบเรื่องรายได้นี้ได้แสดงไว้ในตาราง ก. 3 ของภาคผนวก ก.

### 3.3 ผลการประมาณค่ารายได้ที่ขาดหายไป

ในการวิจัยครั้งนี้ใช้วิธีประมาณค่ารายได้ที่ขาดหายไปในกรณีที่ไม่สนใจกลไกของการตอบเรื่องรายได้ (Ignorable Response Mechanism) เนื่องจากข้อมูลที่นำมาศึกษาครั้งนี้ไม่สามารถหาข้อสันนิษฐานเกี่ยวกับตัวอย่างที่ไม่ให้คำตอบเรื่องรายได้จากแหล่งข้อมูลอื่น ๆ เพื่อนำ

มาทดสอบดูว่า ความน่าจะเป็นของการไม่ตอบรายได้ ขึ้นอยู่กับระดับรายได้หรือไม่และถ้าการตอบรายได้ขึ้นอยู่กับระดับรายได้แล้ว จะมีความสัมพันธ์กันในรูปใด และสิ่งสำคัญประการหนึ่งก็คือ ถึงแม้ว่าจะมีผลการศึกษาของ Greenlees ซึ่งพิจารณาจากครัวเรือนในประเทศสหรัฐอเมริกา แสดงว่า ความน่าจะเป็นของการไม่ตอบรายได้ขึ้นอยู่กับระดับรายได้ และมีความสัมพันธ์เป็นแบบฟังก์ชันลอจิสติก (logistic function) แต่ลักษณะของหน่วยตัวอย่างในแต่ละแหล่ง ไม่น่าจะเหมือนกัน และถ้าจะสรุปว่าลักษณะของหน่วยที่ไม่ตอบเรื่องรายได้ของกรุงเทพมหานครขึ้นอยู่กับระดับรายได้ก็เป็นสิ่งที่ไม่ถูกต้องนัก

ด้วยเหตุผลดังกล่าวข้างต้น วิธีที่เหมาะสมในเวลานี้คือ การประมาณค่ารายได้ที่ขาดหายไป ในกรณีที่ไม่สนใจกลไกของการตอบเรื่องรายได้ ผลการประมาณค่ารายได้ที่ขาดหายไปในแต่ละวิธีเป็นดังนี้

1. วิธีการแบ่งชั้นภูมิเมื่อเลือกตัวอย่างแล้ว เป็นวิธีการประมาณค่าตัวแปรที่ขาดหายไปวิธีหนึ่งที่ยอมรับกันมากในต่างประเทศ เช่น Schaible Brewer Oh และ Scheuren (อ้างถึงใน Greenlees: 1982) Little (1982) ได้นำวิธีนี้ไปวิเคราะห์หาข้อมูลของตัวแปรใด ๆ ที่ขาดหายไป และ Greenlees ได้เสนอว่าควรเป็นวิธีหนึ่งในหลายวิธีที่สามารถนำมาประมาณค่ารายได้ที่ขาดหายไป เพราะวิธีนี้เป็นการหาตัวแปรที่มีความสัมพันธ์อย่างสูงกับตัวแปรรายได้ที่สนใจศึกษา ค่าประมาณที่ได้ ควรจะมีความแม่นยำสูงกว่า การแบ่งชั้นภูมิก่อนเลือกตัวอย่าง สำหรับผลการหาตัวแปรที่มีความสัมพันธ์อย่างสูงกับตัวแปรรายได้เป็นดังนี้

วิธีการหาตัวแปรที่มีความสัมพันธ์อย่างสูงกับตัวแปรรายได้ ได้ใช้ค่าคราเมอร์ (Cramer's V) ซึ่งเป็นค่าสถิติที่แสดงระดับความสัมพันธ์ระหว่างตัวแปร 2 ตัว ค่าคราเมอร์ที่ได้นี้เป็นการปรับค่า  $\chi^2$  ด้วยจำนวนตัวอย่างทั้งหมด และจำนวนชั้นของตัวแปรที่น้อยกว่า ดังนี้

$$V = \left( \frac{\chi^2}{N \cdot \min(r-1, c-1)} \right)^{\frac{1}{2}}$$

โดยที่ V แทน ค่าคราเมอร์ ที่มีค่าตั้งแต่ 0 ถึง 1

r, c แทน จำนวนแถวและสทมภ์ตามลำดับ

N แทน ผลรวมของความถี่ทั้งหมด

ค่าสถิติคราเมอร์ ซึ่งแสดงระดับความสัมพันธ์ระหว่างระดับรายได้หลักของหัวหน้าครัวเรือนกับตัวแปรอื่น ๆ ได้แสดงไว้ในตารางที่ 10

ตารางที่ 10 แสดงค่าคราแมร์รี ของตัวแปรที่มีความสัมพันธ์กับรายได้หลักของหัวหน้าครัวเรือน

ตัวแปร	ค่าคราแมร์รี
ฐานะทาง เศรษฐกิจ	0.3458
รายจ่ายสำหรับครัวเรือน	0.2593
ระดับการศึกษาของหัวหน้าครัวเรือน	0.2115
อาชีพหลักของหัวหน้าครัวเรือน	0.1965
สถานภาพในการทำงานอาชีพหลัก	0.1751
จำนวนชั่วโมงทำงานต่อสัปดาห์	0.1120
เขตที่ครัวเรือนตั้งอยู่	0.1065
ระดับอายุของหัวหน้าครัวเรือน	0.0879
ประสบการณ์ในการทำงานอาชีพหลัก	0.0851

จะเห็นว่า เมื่อพิจารณาจากค่าคราแมร์รี รายได้หลักของหัวหน้าครัวเรือนกับฐานะทางเศรษฐกิจมีระดับความสัมพันธ์สูงกว่าตัวแปรอื่น ๆ ดังนั้น จึงใช้ฐานะทางเศรษฐกิจเป็นตัวแปรแบ่งชั้นภูมิ เมื่อเลือกตัวอย่างแล้ว (Poststratification variable) เนื่องจากตัวแปรฐานะทางเศรษฐกิจ เป็นตัวแปรแบบไม่ต่อเนื่องจึงแบ่งชั้นภูมิออกเป็น 5 ชั้นภูมิอย่างกว้าง ๆ ตามลักษณะของตัวแปรคือ ต่ำมาก ต่ำ ปานกลาง แยก และแยกมาก เป็นชั้นภูมิที่ 1 2 3 4 และ 5 ตามลำดับ

หลังจากแบ่งชั้นภูมิโดยตัวแปรฐานะทางเศรษฐกิจแล้วได้ประมาณค่ารายได้ที่ขาดหายไปดังต่อไปนี้

วิธีที่ 1 การประมาณรายได้ที่ขาดหายไปโดยใช้ค่าเฉลี่ยจากชั้นภูมิ วิธีนี้เป็นการประยุกต์การเดาค่าที่ไม่ทราบโดยใช้ค่าเฉลี่ยของตัวแปรที่สนใจศึกษาที่เราทราบค่า (Beale and Little: 1975) และ Little (1982) ได้เสนอวิธีการประมาณค่าวิธีหนึ่งคือการใช้ค่าเฉลี่ยของตัวแปรที่สนใจศึกษาที่เราทราบค่าจากชั้นภูมิ เมื่อเลือกตัวอย่างแล้วและให้เหตุผลว่าเป็นวิธีการประมาณค่าที่ขาดหายไปที่วิธีหนึ่งในกรณีที่ไม่สนใจกลไกของการตอบ (Ignorable Response Mechanism) เนื่องจากผลการประมาณที่ได้ น่าเชื่อถือได้มากกว่าค่าเฉลี่ยธรรมดาหรือค่าเฉลี่ยจากชั้นภูมิก่อนเลือก



ตัวอย่าง เพราะตัวแปรที่ใช้แบ่งชั้นภูมิ เมื่อ เลือกตัวอย่างแล้วมีความสัมพันธ์กับตัวแปรที่สนใจศึกษา สูงกว่าตัวแปรแบ่งชั้นภูมิ ก่อน เลือกตัวอย่าง

ผลการประมาณค่ารายได้ที่ขาดหายไปโดยใช้ค่าเฉลี่ยของรายได้ที่ทราบค่าจากชั้นภูมิ เมื่อฐานะทาง เศรษฐกิจ เป็นตัวแปรแบ่งชั้นภูมิ ได้ใช้โปรแกรมสำเร็จรูป SPSS ช่วยประมวลผล

ก. สำหรับตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ซึ่ง เลือกมาจากครัว เรือนที่ให้คำตอบ เรื่องรายได้ 1,436 ครัวเรือน โดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัว เรือน จำนวน 61 ครัว เรือน ตามที่ได้อธิบายมาแล้วในหัวข้อ 3.2 จะเหลือหน่วยสำหรับการหาค่าเฉลี่ยจากแต่ละชั้นภูมิ จำนวนรวมทั้งสิ้น 1,375 ครัวเรือน ดังได้แสดงไว้ในตารางที่ 11

ตารางที่ 11 แสดงค่าเฉลี่ยของรายได้ที่ทราบค่าหลังจาก เลือกตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ โดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัว เรือน และมีฐานะทาง เศรษฐกิจ เป็นตัวแปร แบ่งชั้นภูมิ เมื่อ เลือกตัวอย่างแล้ว

ชั้นภูมิที่	ขนาดตัวอย่าง	ค่าเฉลี่ยของรายได้
1	27	14,945
2	273	9,686
3	865	4,069
4	198	2,396
5	12	1,917

นำผลที่ได้จากตารางที่ 11 ไปประมาณค่ารายได้ของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ ทั้ง 61 ครัวเรือนโดยพิจารณาว่าหน่วยใดอยู่ในชั้นภูมิใดจึงประมาณค่ารายได้โดยใช้ค่าเฉลี่ยจาก ชั้นภูมินั้น ตัวอย่างเช่น หน่วยที่ 1 ในตาราง ก. 2 ของภาคผนวก ก. ตกอยู่ในชั้นภูมิที่ 2 (ฐานะดี) ค่าประมาณรายได้จึงมีค่าเท่ากับ 9,686 บาท เป็นต้น สำหรับผลการประมาณค่ารายได้ทั้ง 61 ครัวเรือนได้แสดงไว้ในตารางที่ 13

ข. สำหรับตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ ซึ่งเลือกอย่างสุ่มจำนวน 74 ครัวเรือน ตามที่ได้อธิบายมาแล้วในหัวข้อ 3.2 จะเหลือหน่วยสำหรับการหาค่าเฉลี่ยจากแต่ละชั้นภูมิ จำนวนรวมทั้งสิ้น 1,362 ครัวเรือน ดังได้แสดงไว้ในตารางที่ 12

ตารางที่ 12 แสดงค่าเฉลี่ยของรายได้ที่ทราบค่าหลังจากเลือกตัวแทนอย่างสุ่มเพื่อแทนหน่วยที่ไม่ตอบเรื่องรายได้ และมีฐานะทางเศรษฐกิจเป็นตัวแบ่งชั้นภูมิเมื่อเลือกตัวอย่างแล้ว

ชั้นภูมิที่	ขนาดตัวอย่าง	ค่าเฉลี่ยของรายได้
1	27	14,204
2	271	9,928
3	858	4,099
4	195	2,398
5	11	1,909

นำผลที่ได้จากตารางที่ 12 ไปประมาณค่ารายได้ของตัวแทนหน่วยที่ไม่ตอบเรื่องรายได้ ทั้ง 74 ครัวเรือน โดยพิจารณาว่าหน่วยใดอยู่ในชั้นภูมิใด จึงประมาณค่ารายได้โดยใช้ค่าเฉลี่ยจากชั้นภูมินั้น เช่นเดียวกับที่ทำในกรณีที่เลือกตัวแทนโดยอาศัย ปัจจัยพื้นฐานของหัวหน้าครัวเรือน ดังกล่าวมาแล้วในข้อ ก. สำหรับผลการประมาณค่ารายได้ทั้ง 74 ครัวเรือนได้แสดงไว้ในตารางที่ 14

วิธีที่ 2 การประมาณค่ารายได้ที่ขาดหายไปโดยใช้ค่าสุ่มจากชั้นภูมิ วิธีนี้เป็นที่กล่าวถึงกันมากในต่างประเทศ เพราะเป็นวิธีที่ทำได้ง่ายและสะดวกในทางปฏิบัติถึงแม้ว่าจะไม่ค่อยมีประสิทธิภาพเต็มที่ (Little: 1982) และสิ่งที่น่าสนใจคือน่าจะเป็นวิธีที่เหมาะสมวิธีหนึ่งในการประมาณค่ารายได้ที่ขาดหายไป โดยทั่วไปแล้ววิธีนี้เรียกว่า "hot deck"

ผลการประมาณค่ารายได้ที่ขาดหายไปโดยใช้ค่าสุ่มจากชั้นภูมิ เมื่อฐานะทางเศรษฐกิจเป็นตัวแปรแบ่งชั้นภูมิ ได้ใช้โปรแกรมสำเร็จรูป SPSS สำหรับเลือกตัวอย่างสุ่ม (random sample)

ก. สำหรับตัวแทนหน่วยที่ไม่ตอบในเรื่องราวได้ซึ่งเลือกมาจากครัวเรือนที่ให้คำตอบเรื่องราวได้ 1,436 ครัวเรือนโดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัวเรือน จำนวน 61 ครัวเรือนนั้นจะเหลือหน่วยสำหรับการหาค่าลุ่มจำนวนรวมทั้งสิ้น 1,375 ครัวเรือน และขนาดตัวอย่างย่อยในแต่ละชั้นภูมิได้จากตารางที่ 11 ในการเลือกหน่วยอย่างลุ่มเพื่อประมาณค่ารายได้ของ 61 ครัวเรือนข้างต้นได้ประมาณค่าตามลำดับของ 61 ครัวเรือน เช่นหน่วยที่ 1 ในตาราง ก. 2 ของภาคผนวก ก. ตกอยู่ในชั้นภูมิที่ 2 ก็เลือกตัวอย่างลุ่ม 1 หน่วย จากขนาดตัวอย่างย่อย 273 ครัวเรือนในชั้นภูมิที่ 2 (จากตารางที่ 11) ค่ารายได้ที่ได้จากหน่วยที่เลือกขึ้นมาอย่างลุ่ม จะเป็นค่าประมาณของรายได้ของหน่วยที่ 1 ทำเช่นนี้เรื่อยไปจนครบ 61 ครัวเรือน โดยอาศัยเทคนิคของโปรแกรมสำเร็จรูป SPSS จะได้ตัวอย่างลุ่มที่ไม่มีหน่วยใดซ้ำกันเลย ผลการประมาณค่ารายได้ทั้ง 61 ครัวเรือนได้แสดงไว้ในตารางที่ 13

ข. สำหรับตัวแทนหน่วยที่ไม่ตอบเรื่องราวได้ซึ่งเลือกโดยวิธีลุ่มจำนวน 74 ครัวเรือนจะเหลือหน่วยสำหรับการหาค่าลุ่มจำนวนรวมทั้งสิ้น 1,362 ครัวเรือน ซึ่งเป็นผลที่ได้มาจากหัวข้อ 3.2 และขนาดตัวอย่างย่อยในแต่ละชั้นภูมิได้จากตารางที่ 12 ในการเลือกหน่วยอย่างลุ่มเพื่อประมาณค่ารายได้ของ 74 ครัวเรือนข้างต้นได้ประมาณค่าเช่นเดียวกับข้อ ก. ที่กล่าวมาแล้วข้างต้น ผลการประมาณค่ารายได้ทั้ง 74 ครัวเรือนได้แสดงไว้ในตารางที่ 14

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 13 แสดงค่าจริงและค่าประมาณรายได้ที่ประมาณจากวิธีการแบ่งชั้นภูมิ เมื่อเลือกตัวอย่างแล้ว โดยมีฐานะทางเศรษฐกิจเป็นตัวแปรแบ่งชั้นภูมิ ของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ ซึ่งเลือกขึ้นมาโดยอาศัยปัจจัยของหัวหน้าครัวเรือน

ลำดับที่	ค่าจริง	แบ่งชั้นภูมิโดย ฐานะทาง เศรษฐกิจ	
		ค่าเฉลี่ยจากชั้นภูมิ	ค่าสุ่มจากชั้นภูมิ
1	10,000	9,686	12,000
2	10,000	9,686	25,000
3	2,500	2,396	3,350
4	13,000	9,686	8,000
5	10,000	14,945	10,000
6	25,000	9,686	4,000
7	15,000	14,945	12,000
8	7,100	9,686	8,500
9	10,000	4,069	8,000
10	7,500	4,069	2,150
11	4,000	4,069	2,500
12	1,800	4,069	2,080
13	2,000	2,396	3,000
14	7,500	4,069	3,540
15	3,000	4,069	2,000
16	3,000	2,396	1,560
17	3,000	4,069	18,000
18	4,000	4,069	3,500
19	1,000	4,069	4,500
20	5,000	4,069	1,500
21	580	4,069	3,580

ตารางที่ 13 (ต่อ)

ลำดับที่	ค่าจริง	แบ่งชั้นภูมิโดย ฐานะทางเศรษฐกิจ	
		ค่าเฉลี่ยจากชั้นภูมิ	ค่าส่วนจากชั้นภูมิ
22	2,800	9,686	6,000
23	7,000	2,396	3,400
24	25,000	9,686	10,000
25	3,000	4,069	5,500
26	6,000	4,069	4,830
27	6,000	9,686	15,000
28	20,000	4,069	3,530
29	3,000	4,069	3,800
30	3,500	4,069	15,000
31	1,000	4,069	1,200
32	2,500	4,069	2,400
33	4,000	4,069	4,500
34	25,000	9,686	5,000
35	3,000	4,069	3,800
36	1,500	4,069	1,000
37	2,500	4,069	2,000
38	1,800	4,069	4,700
39	9,000	4,069	800
40	3,000	4,069	2,000
41	4,000	4,069	3,500
42	3,200	4,069	2,000

ตารางที่ 13 (ต่อ)

ลำดับที่	ค่าจริง	แบ่งชั้นภูมิโดย ฐานะทางเศรษฐกิจ	
		ค่าเฉลี่ยจากชั้นภูมิ	ค่าสูงจากชั้นภูมิ
43	20,000	4,069	6,000
44	9,000	4,069	11,000
45	13,500	9,686	3,500
46	30,000	9,686	3,100
47	2,910	9,686	5,000
48	2,400	4,069	4,440
49	4,000	4,069	1,500
50	9,000	4,069	3,500
51	5,000	2,396	3,000
52	3,000	4,069	3,000
53	1,500	4,069	3,000
54	2,500	4,069	2,700
55	1,500	2,396	3,000
56	750	4,069	3,000
57	500	2,396	2,000
58	6,500	9,686	4,000
59	1,560	2,396	3,500
60	3,000	2,396	600
61	80	2,396	3,000

ตารางที่ 14 แสดงค่าจริงและค่าประมาณรายได้ที่ประมาณจากวิธีการแบ่งชั้นภูมิ เมื่อเลือกตัวอย่างแล้ว โดยมีฐานะทาง เศรษฐกิจ เป็นตัวแปรแบ่งชั้นภูมิของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ซึ่ง เลือกขึ้นมาโดยสุ่ม

ลำดับที่	ค่าจริง	แบ่งชั้นภูมิโดย ฐานะทาง เศรษฐกิจ	
		ค่าเฉลี่ยจากชั้นภูมิ	ค่าสุ่มจากชั้นภูมิ
1	3,000	4,099	3,500
2	3,500	4,099	4,000
3	2,500	4,099	2,000
4	9,000	2,398	1,800
5	34,500	9,928	10,000
6	300	4,099	2,000
7	500	2,398	3,000
8	2,000	4,099	580
9	4,000	4,099	4,600
10	3,000	9,928	10,000
11	35,000	14,204	45,000
12	31,000	9,928	6,500
13	1,500	2,398	2,010
14	2,110	2,398	4,600
15	3,000	4,099	600
16	800	4,099	3,500
17	900	4,099	3,000
18	3,540	4,099	9,000
19	12,000	4,099	5,000
20	2,000	1,909	3,000
21	3,500	4,099	5,000
22	2,000	4,099	3,000
23	4,900	4,099	1,890
24	2,600	2,398	450

ตารางที่ 14 (ต่อ)

ลำดับที่	ค่าจริง	แบ่งชั้นภูมิโดย ฐานะทาง เศรษฐกิจ	
		ค่าเฉลี่ยจากชั้นภูมิ	ค่าสุ่มจากชั้นภูมิ
25	13,000	9,928	25,000
26	5,000	4,099	7,610
27	1,350	9,928	9,000
28	2,000	4,099	5,000
29	2,500	4,099	1,500
30	2,000	2,398	1,500
31	6,200	9,928	20,000
32	4,200	9,928	3,100
33	2,000	4,099	2,100
34	1,000	4,099	3,000
35	4,600	9,928	10,000
36	15,000	4,099	3,000
37	2,000	2,398	1,500
38	1,500	4,099	2,200
39	20,000	4,099	7,000
40	4,300	9,928	5,000
41	6,000	4,099	2,000
42	8,000	4,099	5,000
43	3,000	4,099	4,000
44	1,000	2,398	1,500
45	1,000	2,398	2,000
46	3,000	4,099	1,660
47	2,500	4,099	6,500
48	2,000	2,398	3,000
49	4,000	4,099	4,000



ตารางที่ 14 (ต่อ)

ลำดับที่	ค่าจริง	แบ่งชั้นภูมิโดย ฐานะทาง เศรษฐกิจ	
		ค่าเฉลี่ยจากชั้นภูมิ	ค่าสูงจากชั้นภูมิ
50	5,000	9,928	6,000
51	2,100	4,099	3,000
52	1,260	2,398	3,000
53	3,250	9,928	3,000
54	6,000	4,099	3,000
55	1,000	14,204	12,000
56	2,000	4,099	1,900
57	3,300	4,099	3,000
58	1,600	4,099	1,500
59	6,500	9,928	4,000
60	4,000	9,928	5,000
61	6,000	4,099	2,000
62	7,000	4,099	7,000
63	120	9,928	3,000
64	5,200	4,099	7,500
65	1,000	4,099	580
66	2,800	4,099	550
67	5,000	2,398	1,800
68	3,000	4,099	4,000
69	2,540	4,099	4,000
70	830	4,099	2,500
71	2,000	4,099	4,000
72	2,000	4,099	6,500
73	4,500	4,099	3,000
74	3,000	2,398	3,000

## 2. วิธีประมาณค่ารายได้ที่ขาดหายไปจากสมการถดถอยพหุเชิงเส้น

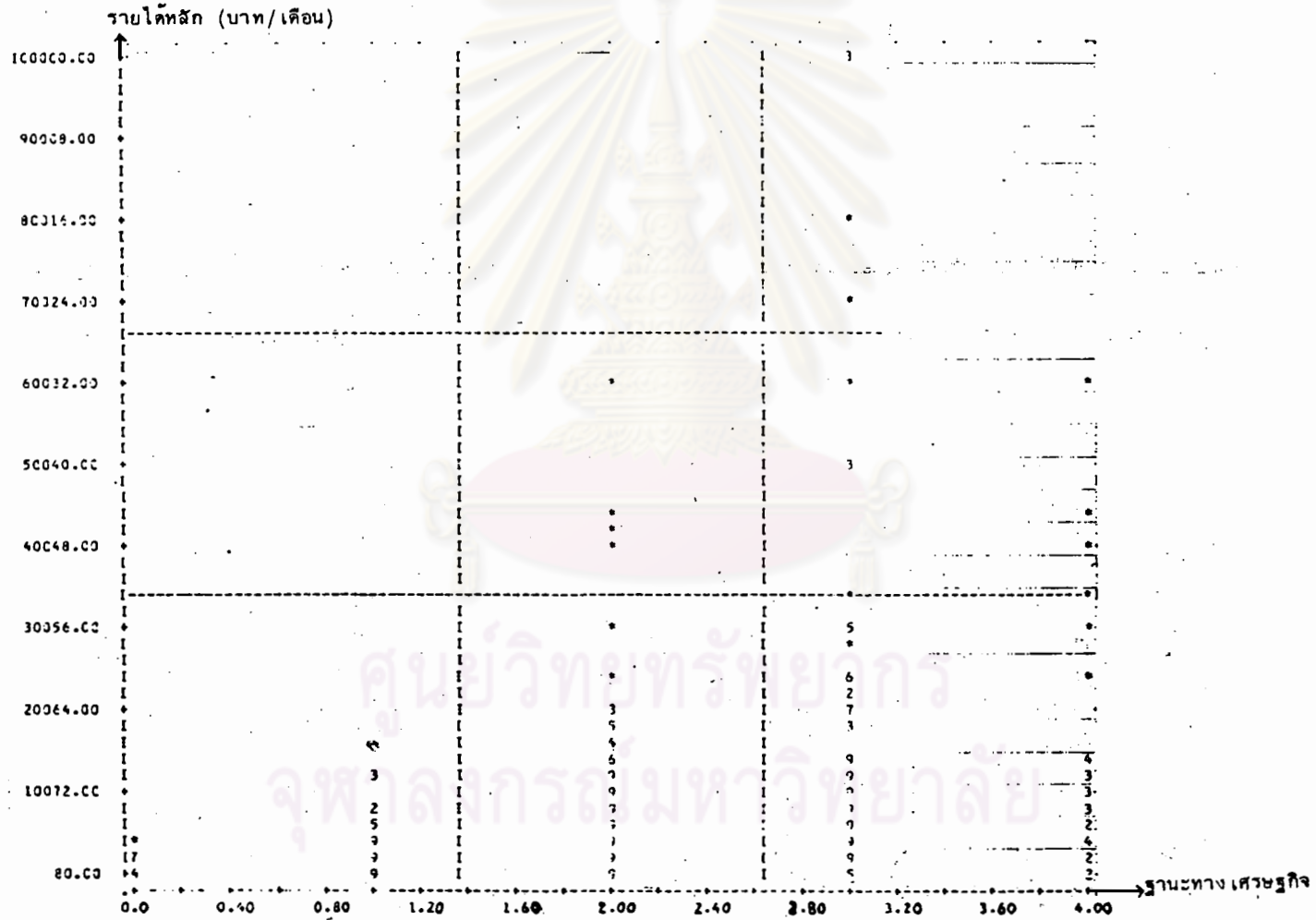
ในกรณีที่ไม่สนใจว่าความน่าจะเป็นของการตอบขึ้นอยู่กับระดับรายได้แล้ววิธีที่เหมาะสมในการประมาณค่ารายได้ที่ขาดหายไป วิธีหนึ่งคือ การประมาณค่าที่ได้จากสมการถดถอยพหุแบบเชิงเส้น (Greenlees: 1982) ส่วนการประมาณค่าพารามิเตอร์ที่เป็นตัวกำหนดค่ารายได้นั้น จะใช้ตัวอย่างเฉพาะที่ตอบเรื่องรายได้เท่านั้น และจะไม่นำหน่วยที่สมมติว่า ขาดหายไปจากหัวข้อ 3.2 มาคิด เมื่อได้สมการพยากรณ์แล้ว จึงจะประมาณค่ารายได้ที่ขาดหายไปโดยอาศัยสมการพยากรณ์ที่ได้

สำหรับการหาตัวกำหนดรายได้อันดับของหัวหน้าครัวเรือนนั้น ได้คำนึงถึงปัจจัยที่มีระดับความสัมพันธ์กับรายได้อย่างสูงและรองลงมาดังได้แสดงมาแล้วในตารางที่ 10 และปัจจัยที่ใช้ในการวิเคราะห์สมการถดถอยได้มีแนวความคิดจากผู้ที่เคยวิเคราะห์เกี่ยวกับตัวกำหนดรายได้มาแล้ว

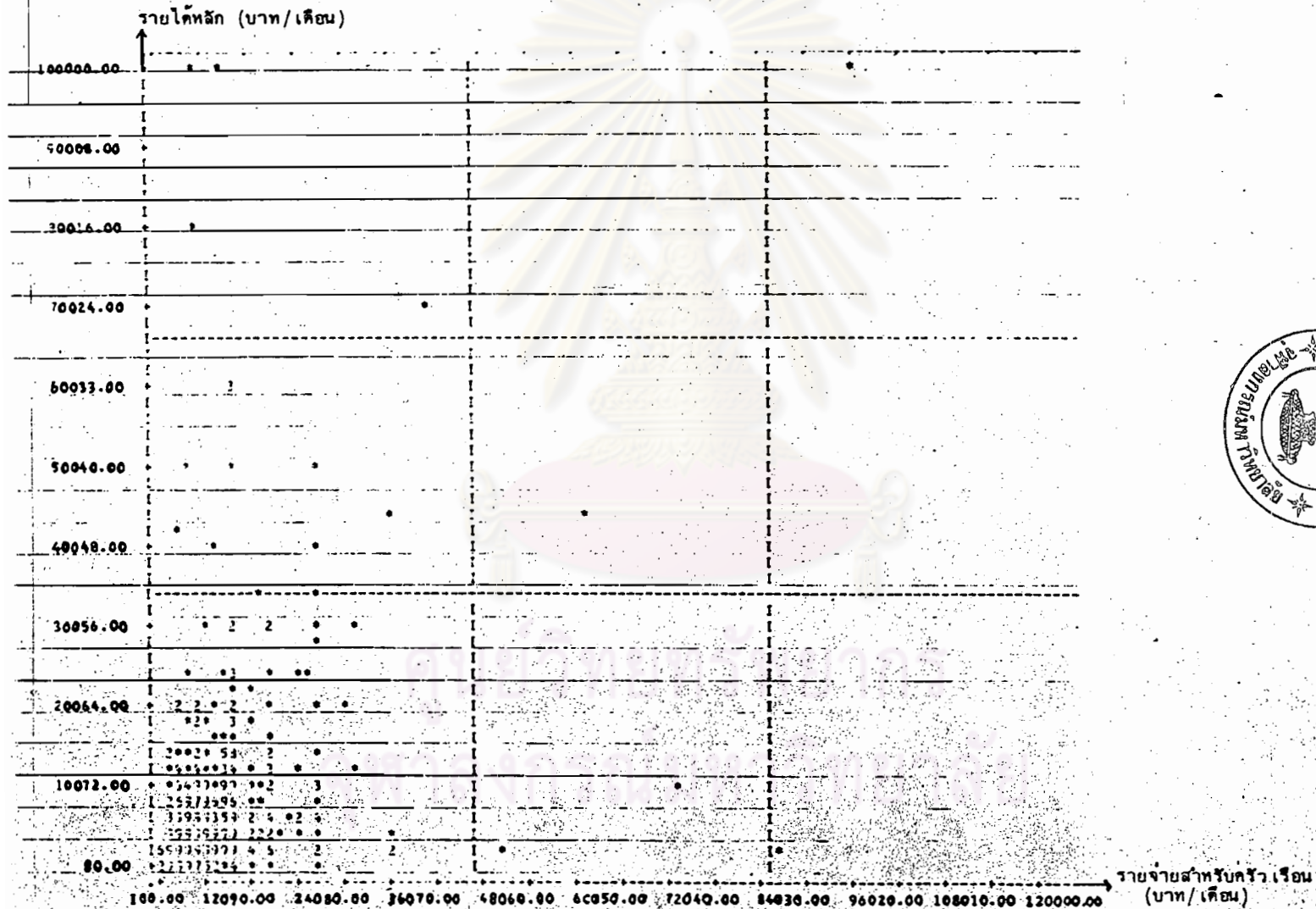
Greenlees and others (1982) ได้ศึกษามาแล้วว่า จากตัวแปรทั่วไปของสมการถดถอยพหุเชิงเส้น  $Y_i = \alpha + \beta X_i + \epsilon_i$  เป็นฟังก์ชัน semi-logarithm ของรายได้ คือตัวแปรตาม Y อยู่ในรูปลอการิทึมของรายได้ ซึ่งในทางปฏิบัติแล้วเราไม่สามารถทดสอบสมมติฐานเกี่ยวกับตัว  $\epsilon$  ได้ว่า มีการแจกแจงแบบปกติโดยที่มีค่าเฉลี่ยเป็นศูนย์และความแปรปรวนเป็น  $\sigma^2$  ด้วยเหตุผลสำคัญประการหนึ่งคือ ไม่สามารถหาค่าสังเกตที่ไม่ตอบของตัวแปรตาม Y และ Greenlees ได้อาศัยแนวความคิดมาจากนักสถิติอื่น ๆ ที่เคยศึกษาวิจัยมาแล้ว ประกอบกับเขาสามารถวิเคราะห์กับข้อมูลที่ ใช้ศึกษา เนื่องจาก สามารถหาแหล่งข้อมูลเกี่ยวกับหน่วยตัวอย่างที่ต้องการศึกษา ตัวแปรดังกล่าวนี้ได้

ก่อนที่จะนำตัวแปรทั่วไปของสมการถดถอยพหุเชิงเส้น โดยที่ตัวแปรตาม Y อยู่ในรูปลอการิทึมของรายได้มาใช้ในการวิเคราะห์นั้น ได้ศึกษาเปรียบเทียบระหว่างตัวแปรรายได้ กับตัวแปรลอการิทึมของรายได้ ว่าตัวแปรใดมีความสัมพันธ์ในรูปเชิงเส้น กับปัจจัยอื่น ๆ ที่จะนำมาเป็นตัวกำหนดรายได้ ได้ดีกว่ากัน ในที่นี้ได้ดูความสัมพันธ์ระหว่างตัวแปรรายได้กับตัวแปรฐานะทางเศรษฐกิจ และตัวแปรรายจ่ายสำหรับครัวเรือน เนื่องจากตัวแปรทั้งสองนี้วัดระดับความสัมพันธ์กับตัวแปรรายได้ได้สูงกว่าตัวแปรอื่น ๆ (จากตารางที่ 10) และดูความสัมพันธ์ระหว่างตัวแปรลอการิทึมฐานสิบของรายได้ กับตัวแปรฐานะทางเศรษฐกิจและตัวแปร รายจ่ายสำหรับครัวเรือน เช่นเดียวกัน การดูความสัมพันธ์ระหว่างตัวแปรแต่ละคู่ดังกล่าวมาแล้ว ได้พิจารณาจากแผนภาพการกระจาย (scatter diagram) ดังภาพที่ 1 ถึงภาพที่ 4

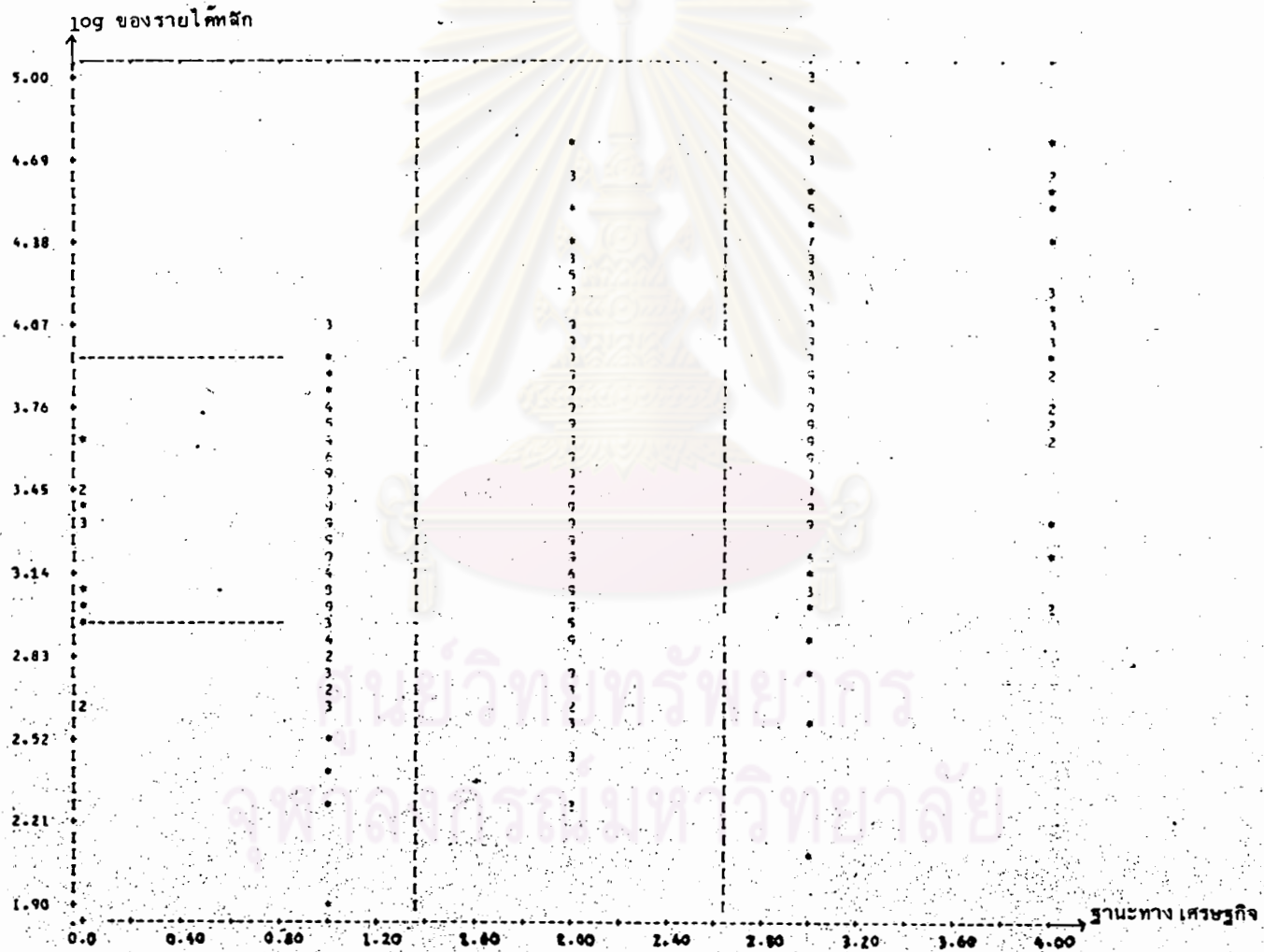
ภาพที่ 1 แสดงแผนภาพการกระจายระหว่างรายได้หลักของหัวหน้าครัวเรือนและฐานะทางเศรษฐกิจของผู้ให้สัมภาษณ์



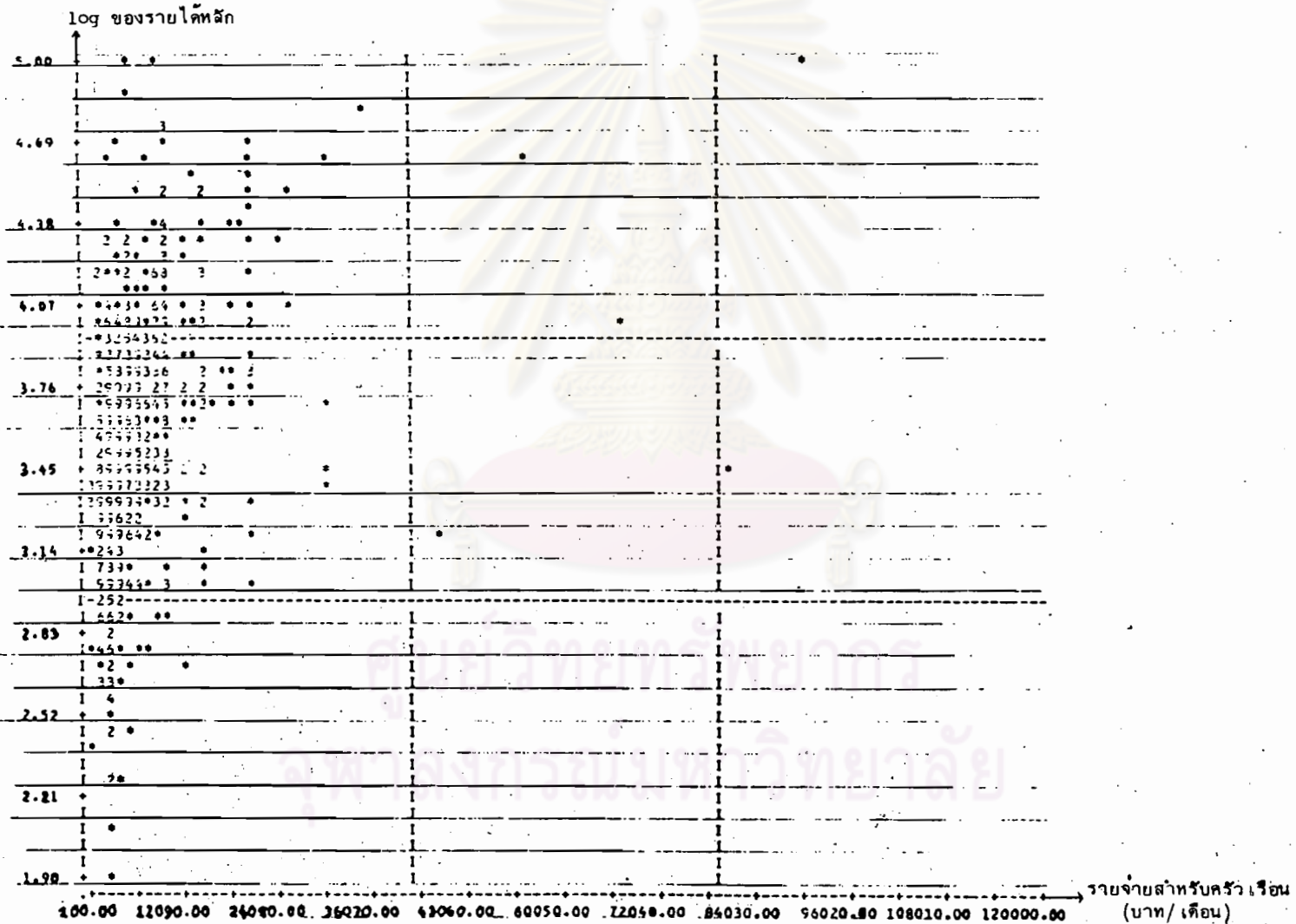
ภาพที่ 2 แสดงแผนภาพการกระจาย ระหว่างรายได้หลักของหัวหน้าครัวเรือนและรายจ่ายสำหรับครัวเรือน



ภาพที่ 3 แสดงแผนภาพการกระจายระหว่างค่า log ของรายได้อีกของหัวหน้าครัวเรือนและฐานะทางเศรษฐกิจของผู้ให้สัมภาษณ์



ภาพที่ 4 แสดงแผนภาพการกระจายระหว่างค่า log ของรายได้หลักของหัวหน้าครัวเรือนและรายจ่ายสำหรับครัวเรือน



จากภาพที่ 1 จะเห็นว่าความสัมพันธ์ระหว่างรายได้หลักของหัวหน้าครัวเรือนและฐานะทางเศรษฐกิจ มีค่ากระจายออกไปมาก เช่นเดียวกับ รูปที่ 2 ซึ่งแสดง ความสัมพันธ์ระหว่างรายได้หลักของหัวหน้าครัวเรือน และรายจ่ายสำหรับครัวเรือน ส่วนภาพที่ 3 จะเห็นว่าความสัมพันธ์ระหว่างลอการิทึมฐานสิบของรายได้หลักของหัวหน้าครัวเรือนกับฐานะทางเศรษฐกิจ มีค่าลำดับที่กระจายออกจากกลุ่มน้อยลง เช่นเดียวกับภาพที่ 4 ที่แสดงความสัมพันธ์ระหว่างลอการิทึมฐานสิบของรายได้หลักของหัวหน้าครัวเรือน และรายจ่ายสำหรับครัวเรือน ประกอบกับแนวความคิดของ Gronau (1973) พอสรุปได้ว่า "ลักษณะการแจกแจงของรายได้ของประชากรมีลักษณะเบ้ขวา ถ้ามีการแปลงข้อมูลให้อยู่ในรูปลอการิทึมฐานสิบแล้วการกระจายจะแคบลงจนเข้าสู่การแจกแจงแบบปกติ" ในการวิเคราะห์ความถดถอยครั้งนี้จึงใช้ตัวแบบทั่วไปของการถดถอยพหุเชิงเส้นคือ  $Y_i = X_i\beta + \epsilon_i$  โดยที่  $Y_i$  เป็นค่าลอการิทึมฐานสิบของรายได้ และ  $\epsilon_i$  อยู่ในรูปลอการิทึมฐานสิบที่มีการแจกแจงแบบปกติด้วยค่าเฉลี่ยศูนย์และความแปรปรวน  $\sigma^2$  รูปแบบทั่วไปนี้จะพบได้ทั่วไปในการวิเคราะห์ความถดถอยเกี่ยวกับค่าตอบแทน เช่น Mincer (1974: อ้างถึงใน Greenlees and others 1982: 256) Gronau (1973) Greenlees and others (1982)

สำหรับการวิเคราะห์ความถดถอยครั้งนี้ใช้กับขนาดตัวอย่าง 3 ชุดคือ

ข้อมูลชุดที่ 1 เป็นชุดข้อมูลที่มีหน่วยที่ตอบรายได้ ครบทุกหน่วยก่อนที่จะมีการเลือกตัวแทนหน่วยที่ไม่ตอบเรื่องรายได้ คือ มีจำนวนทั้งหมด 1,436 ครัวเรือน ในการวิเคราะห์ขั้นนี้เพื่อศึกษาดูว่ามีปัจจัยอะไรบ้างที่เป็นตัวกำหนดในการตอบรายได้ ทั้ง 1,436 ครัวเรือน

ข้อมูลชุดที่ 2 เป็นข้อมูลชุดที่ไม่รวมตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ ซึ่งเลือกมาจากครัวเรือนที่ให้คำตอบเรื่องรายได้ 1,436 ครัวเรือนโดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัวเรือนจำนวน 61 ครัวเรือนดังได้กล่าวมาแล้ว ดังนั้น จะเหลือหน่วยสำหรับการวิเคราะห์ความถดถอยจำนวน 1,375 ครัวเรือนเพื่อประมาณค่ารายได้ของ 61 ครัวเรือนในการประมาณค่านี้อาศัยค่าประมาณสัมประสิทธิ์ความถดถอย และตัวแปรอิสระในแต่ละหน่วยของตัวแปรทั้ง 61 ครัวเรือน นอกจากนี้ยังศึกษาดูว่าถ้าประมาณข้อมูลชุดที่ 2 นี้ด้วยสมการถดถอยแล้วผลการประมาณจะใช้ได้หรือไม่

ข้อมูลชุดที่ 3 เป็นข้อมูลชุดที่ไม่รวมตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ ซึ่งเลือกมาจากครัวเรือนที่ให้คำตอบเรื่องรายได้ 1,436 ครัวเรือนโดยวิธีสุ่ม จำนวน 74 ครัวเรือน ดังนั้นจะเหลือหน่วยสำหรับการวิเคราะห์ความถดถอยจำนวน 1,362 ครัวเรือน ในการประมาณค่ารายได้ของตัวแทนทั้ง 74 ครัวเรือน ได้ทำเช่นเดียวกับข้อมูลชุดที่ 2 และนำผลการประมาณที่ได้มีไปศึกษาดูว่าวิธีการประมาณด้วยสมการถดถอยสำหรับข้อมูลชุดที่ 3 นี้ซึ่ง เป็นการสมมติว่าข้อมูลรายได้ขาดหายไปโดยสุ่มจะใช้ได้ผลหรือไม่

สำหรับตัวแบบขั้นต้นที่ใช้กำหนดตัวแปรรายได้หลักของหัวหน้าครัวเรือนในการวิเคราะห์  
ความถดถอย เป็นดังนี้

$$\begin{aligned} \text{logincome} = & \beta_0 + \beta_1 (\text{ECONS}) + \beta_2 (\text{EXPEN}) + \beta_3 (\text{EDUC}) + \\ & \beta_4 (\text{PROF}) + \beta_5 (\text{ADMI}) + \beta_6 (\text{CLER}) + \beta_7 (\text{SALE}) + \\ & \beta_8 (\text{FARM}) + \beta_9 (\text{MINE}) + \beta_{10} (\text{TRAN}) + \beta_{11} (\text{CRAF}) + \\ & \beta_{12} (\text{STATUS0}) + \beta_{13} (\text{STATUS1}) + \beta_{14} (\text{STATUS2}) + \\ & \beta_{15} (\text{STATUS3}) + \beta_{16} (\text{STATUS4}) + \beta_{17} (\text{HRSPWK}) + \\ & \beta_{18} (\text{AGE}) + \beta_{19} (\text{EXPR}) + \beta_{20} (\text{REG 1}) + \beta_{21} (\text{REG 2}) + \varepsilon \end{aligned}$$

- เมื่อ  $\beta_0$  = ค่าคงที่  
 $\beta$  = สัมประสิทธิ์ความถดถอยของแต่ละตัวแปรอิสระ  
 $\varepsilon$  = เป็นค่าความคลาดเคลื่อน =  $\log E \sim N(0, \sigma^2)$

โดยที่ตัวแปรอิสระได้นิยามไว้ใน ภาคผนวก ข

ผลการวิเคราะห์ความถดถอยเพื่อประมาณค่าพารามิเตอร์ โดยใช้วิธี Stepwise Multiple Regression ที่ได้จากโปรแกรมสำเร็จรูป SPSS ของข้อมูลทั้ง 3 ชุด ดังกล่าวแล้วข้างต้น ได้ค่าประมาณสัมประสิทธิ์ความถดถอย ( $\hat{\beta}$ ) ค่าความคลาดเคลื่อนมาตรฐาน (Standard error) ดังตาราง  
ที่ 15

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ตารางที่ 15 แสดงการประมาณค่าพารามิเตอร์และค่าความคลาดเคลื่อนมาตรฐานที่ได้จาก  
สมการถดถอย

ตัวแปรอิสระ	ค่าประมาณสัมประสิทธิ์ความถดถอย		
	ข้อมูลชุดที่ 1	ข้อมูลชุดที่ 2	ข้อมูลชุดที่ 3
ECONS	0.14857 (0.01416)	0.14333 (0.144)	0.15708 (0.01446)
EXPEN	0.00001 (0.00000)	0.00001 (0.00000)	0.00001 (0.00000)
FARM	-0.15228 (0.04472)	-0.1774 (0.04592)	-0.22227 (0.03673)
CRAF	0.0306 (0.03349)	0.02101 (0.03357)	-0.04546 (0.0208)
HRSPWK	0.0018 (0.00042)	0.00184 (0.00043)	0.0018 (0.00042)
EDUC	0.016 (0.0029)	0.0161 (0.00293)	0.01835 (0.00254)
EXPR	0.00345 (0.00081)	0.0035 (0.00083)	0.00424 (0.00083)
STATUS0	0.10579 (0.0216)	0.09737 (0.02213)	0.10638 (0.02079)
STATUS4	0.1339 (0.03459)	0.13462 (0.03458)	0.13781 (0.03481)
ADMI	0.2379 (0.06248)	0.24076 (0.06231)	0.12375 (0.05559)
TRAN	0.1962 (0.05272)	0.19056 (0.05257)	0.11203 (0.04643)
SALE	0.0898 (0.03382)	0.08042 (0.03395)	-
PROF	0.1168 (0.04529)	0.11824 (0.04534)	-
CLER	0.09527 (0.04112)	0.09123 (0.04108)	-
ค่าคงที่	2.76086	2.77290	2.786509
Standard error	0.30296	0.30157	0.30373
R	0.58865	0.58979	0.58668
จำนวนตัวอย่าง	1,436	1,375	1,362

หมายเหตุ ตัวเลขในวงเล็บเป็นค่าความคลาดเคลื่อนมาตรฐาน

จากตารางที่ 15 เมื่อกำหนดตัวประมาณ  $\hat{\beta}$  ด้วยวิธี Stepwise Multiple Regression หลักเกณฑ์ในการตัดตัวแปรอิสระใด ๆ ออกจากสมการในแต่ละขั้นของวิธีนี้พิจารณาจากค่า Partial - F ของตัวแปรอิสระที่มีอยู่ในสมการของขั้นนั้น ๆ เปรียบเทียบกับค่า  $F_{\alpha}(1, n-p)$  ( $p$  หมายถึง จำนวนตัวแปรอิสระที่มีอยู่ทั้งหมดในสมการของขั้นนั้น ๆ) ถ้าค่า Partial - F ของตัวแปรอิสระใดมากกว่า  $F_{\alpha}(1, n-p)$  ก็รวมเอาตัวแปรอิสระตัวนั้นไว้ในสมการ แต่ถ้าค่า Partial - F ใดน้อยกว่า  $F_{\alpha}(1, n-p)$  ก็จะตัดตัวแปรอิสระนั้นออกจากสมการของขั้นนั้น ๆ จะพิจารณาเช่นนี้ ทุกครั้งที่มีตัวแปรอิสระเพิ่ม เข้ามาในสมการใหม่ และสิ้นสุดเมื่อไม่มีตัวแปรอิสระใดที่จะเข้ามาอยู่ในสมการ สมการที่ได้จะมีตัวแปรอิสระบางตัวที่เหมาะสมอยู่ในสมการ กล่าวคือ มีตัวแปรอิสระทั้งหมด 19 ตัว ผลลัพธ์สุดท้ายที่จะเหลือตัวแปรอิสระในสมการเป็นดังนี้

1) จากข้อมูลชุดที่ 1 ซึ่งเป็นชุดข้อมูลที่มีหน่วยตอบรายได้ ทั้งหมด 1,436 ครอบครัว หลังจากใช้วิธี Stepwise Multiple Regression แล้ว เหลือตัวแปรอิสระในสมการทั้งหมด 14 ตัว ผลการประมาณพารามิเตอร์  $\beta$  จากสมการที่ 2 ในตารางที่ 15 ชี้ให้เห็นว่า ผลกระทบของฐานะทางเศรษฐกิจต่อค่าลอการิทึมของรายได้มีค่าเป็นบวก เช่นเดียวกับรายจ่ายของครัวเรือน กลุ่มอาชีพที่มีรายได้มากที่สุดคือ การบริหารและการจัดการ จำนวนชั่วโมงมีความสัมพันธ์กับรายได้เป็นบวก เช่นเดียวกับการศึกษาและประสบการณ์ในการทำงาน สำหรับสถานภาพในการทำงานอาชีพหลัก การเป็นลูกจ้าง/พนักงานรัฐวิสาหกิจ มีรายได้สูงกว่าเป็นเจ้าของกิจการ กล่าวโดยสรุปแล้ว การมีฐานะทางเศรษฐกิจสูงขึ้น การใช้จ่ายสำหรับครัวเรือนเพิ่มขึ้น มีชั่วโมงทำงานเพิ่มขึ้น ระดับการศึกษาสูงขึ้น มีประสบการณ์ในการทำงานมากขึ้น ย่อมมีรายได้เพิ่มขึ้น โดยเฉพาะอย่างยิ่งทำงานระดับบริหารและเป็นพนักงานรัฐวิสาหกิจ

2) จากข้อมูลชุดที่ 2 เป็นชุดข้อมูลที่ไม่รวมตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ซึ่งเลือกขึ้นมาโดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัวเรือนจำนวน 61 ครัวเรือน ดังได้กล่าวมาแล้วจะเหลือหน่วยสำหรับการวิเคราะห์ความถดถอย 1,375 ครัวเรือน หลังจากใช้วิธี Stepwise Multiple Regression แล้ว เหลือตัวแปรอิสระในสมการทั้งหมด 14 ตัว และผลการประมาณค่าพารามิเตอร์ จากสมการที่ 3 ในตารางที่ 15 ให้ค่าใกล้เคียงกับการใช้หน่วยทุกหน่วยในการวิเคราะห์คือ 1,436 ครัวเรือน ดังนั้นพอสรุปได้ว่า สมการถดถอยที่พยากรณ์ได้นี้ น่าจะเป็นสมการพยากรณ์ที่ดีสำหรับการประมาณค่ารายได้ที่ทำให้ขาดหายไปทั้ง 61 ครัวเรือน

ผลการประมาณค่ารายได้ของตัวแทนทั้ง 61 ครว เรือนนี้ ได้ใช้ตัวประมาณ  $\hat{\beta}$  จากสมการถดถอยที่พยากรณ์ได้ในสดมภ์ที่ 3 ของตารางที่ 15 และนำค่าตัวแปรอิสระที่ได้จาก 61 ครว เรือน ซึ่งแสดงไว้ในตาราง ก.2 ของภาคผนวก ก, ไปแทนในสมการดังกล่าวแล้วจะได้ค่าประมาณของลอการิทึมฐานสิบของรายได้ ทั้ง 61 ครว เรือน จากนั้นนำผลลัพธ์ที่ได้ไปหาค่าแอนติลอการิทึม (antilogarithm) จะเป็นค่าประมาณของรายได้ที่ต้องการ ดังได้แสดงไว้ในตารางที่ 16

ตารางที่ 16 แสดงค่าจริงและค่าประมาณรายได้ที่ประมาณจากสมการถดถอยของตัวแทนหน่วยที่ไม่ตอบ เรืองรายได้ ซึ่งเลือกขึ้นมาโดยอาศัยปัจจัยของหัวหน้าครว เรือน

ลำดับที่	ค่าจริง	ค่าที่ประมาณได้จากสมการถดถอย
1	10,000	4,611
2	10,000	5,379
3	2,500	1,576
4	13,000	5,640
5	10,000	9,431
6	25,000	5,307
7	15,000	10,872
8	7,100	6,216
9	10,000	2,723
10	7,500	2,976
11	4,000	2,789
12	1,800	2,115
13	2,000	2,663
14	7,500	3,623
15	3,000	3,161
16	3,000	1,996
17	3,000	3,129

## ตารางที่ 16 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่ประมาณได้จากสมการถดถอย
18	4,000	2,816
19	1,000	2,683
20	5,000	2,092
21	580	1,444
22	2,800	5,492
23	7,000	2,207
24	25,000	5,898
25	3,000	3,243
26	6,000	3,864
27	6,000	5,092
28	20,000	3,660
29	3,000	3,308
30	3,500	2,141
31	1,000	2,190
32	2,500	4,184
33	4,000	3,826
34	25,000	4,598
35	3,000	3,754
36	1,500	2,573
37	2,500	3,488
38	1,800	3,314
39	9,000	4,320
40	3,000	3,440
41	4,000	1,619
42	3,200	2,337
43	20,000	1,735

ตารางที่ 16 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่ประมาณได้จากสมการถดถอย
44	9,000	4,218
45	13,500	5,720
46	30,000	9,944
47	2,910	2,363
48	2,400	2,516
49	4,000	3,697
50	9,000	5,063
51	5,000	3,498
52	3,000	3,757
53	1,500	2,076
54	2,500	1,939
55	1,500	2,381
56	750	1,820
57	500	1,193
58	6,500	3,694
59	1,560	1,389
60	3,000	3,179
61	80	1,645

3) จากข้อมูลชุดที่ 3 เป็นชุดข้อมูลที่ไม่รวมตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ ซึ่งเลือกขึ้นมาโดยวิธีสุ่ม จำนวน 74 ครว เรือน ดังได้กล่าวมาแล้ว หลังจากใช้วิธี Stepwise Multiple Regression โดยมีหน่วยตัวอย่างสำหรับการวิเคราะห์ 1,362 ครว เรือน คงเหลือตัวแปรอิสระในสมการทั้งหมดเพียง 11 ตัวเท่านั้น คือกลุ่มอาชีพค้าขาย วิชาชีพ และ เสมียนพนักงานไม่เข้าไปอยู่ในสมการ ถ้าดูผลการประมาณค่าพารามิเตอร์  $\beta$  และ  $\sigma$  จากสมการที่ 4 ในตารางที่ 15 ค่าที่ได้ไม่แตกต่างกันมากนักส่วนใหญ่มักจะต่างกันในแต่ละทศนิยมที่ 3 ถึงทศนิยมที่ 4 มีเพียงบางตัวเท่านั้นที่ต่างกันในแต่ละทศนิยมที่ 1 และเป็นกลุ่มตัวแปรดัมมี่ เช่น กลุ่มอาชีพ ผลการประมาณรายได้จึงไม่น่าต่างกันมากนัก

ผลการประมาณค่ารายได้ของตัวแทนทั้ง 74 ครว เรือนได้ใช้ตัวประมาณ  $\hat{\beta}$  จากสมการถดถอยที่พยากรณ์ได้ในสมการที่ 4 ของตารางที่ 15 และนำค่าตัวแปรอิสระที่ได้จาก 74 ครว เรือน ซึ่งแสดงไว้ในตาราง ก. 3 ของภาคผนวก ก. ไปแทนในสมการดังกล่าวแล้ว จะได้ค่าประมาณของลอการิทึมฐานสิบของรายได้ ของทั้ง 74 ครว เรือน จากนั้นนำผลลัพธ์ที่ได้นี้ไปหาค่าแอนตี้ลอการิทึมจะเป็นค่าประมาณของรายได้ที่ต้องการดังได้แสดงไว้ในตารางที่ 17

ตารางที่ 17 แสดงค่าจริงและค่าประมาณรายได้ที่ประมาณจากสมการถดถอยของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ ซึ่ง เลือกขึ้นมาโดยสุ่ม

ลำดับที่	ค่าจริง	ค่าที่ประมาณได้จากสมการถดถอย
1	3,000	2,261
2	3,500	3,196
3	2,500	4,424
4	9,000	2,718
5	34,500	7,735
6	300	1,645
7	500	1,637
8	2,000	2,371
9	4,000	3,915
10	3,000	3,119

## ตารางที่ 17 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่ประมาณได้จากสมการถดถอย
11	35,000	11,621
12	31,000	7,993
13	1,500	939
14	2,110	1,757
15	3,000	2,980
16	800	3,906
17	900	2,175
18	3,540	3,691
19	12,000	4,092
20	2,000	1,176
21	3,500	2,727
22	2,000	3,636
23	4,900	3,780
24	2,600	1,531
25	13,000	5,763
26	5,000	2,565
27	1,350	6,388
28	2,000	3,421
29	2,500	3,565
30	2,000	1,598
31	6,200	9,556
32	4,200	5,088
33	2,000	2,617
34	1,000	3,399
35	4,600	8,243
36	15,000	3,443

## ตารางที่ 17 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่ประมาณได้จากสมการถดถอย
37	2,000	1,600
38	1,500	2,601
39	20,000	3,653
40	4,300	5,723
41	6,000	4,928
42	8,000	3,589
43	3,000	2,780
44	1,000	1,768
45	1,000	2,156
46	3,000	3,243
47	2,500	2,924
48	2,000	1,536
49	4,000	3,296
50	5,000	5,858
51	2,100	3,964
52	1,260	1,482
53	3,250	5,670
54	6,000	3,333
55	1,000	12,670
56	2,000	3,361
57	3,300	2,837
58	1,600	1,417
59	6,500	5,484
60	4,000	7,210
61	6,000	4,647
62	7,000	4,595



## ตารางที่ 17 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่ประมาณได้จากสมการถดถอย
63	120	2,677
64	5,200	4,266
65	1,000	2,819
66	2,800	3,722
67	5,000	3,416
68	3,000	3,194
69	2,540	2,790
70	830	1,957
71	2,000	3,652
72	2,000	3,186
73	4,500	3,388
74	3,000	3,155

## 3.4 การเปรียบเทียบค่าเฉลี่ยของร้อยละของผลต่างระหว่างค่าประมาณกับค่าจริงของรายได้

เป็นการศึกษาดูว่า โดยเฉลี่ยแล้ว วิธีการประมาณค่ารายได้ที่ขาดหายไป ที่ได้ศึกษามา 3 วิธีคือ วิธีประมาณโดยใช้ค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิ ซึ่งทำภายหลังจากแบ่งชั้นภูมิเมื่อเลือกตัวอย่างแล้ว และวิธีประมาณโดยใช้สมการถดถอยวิธีใดจะให้ค่าความแตกต่างจากค่าจริงน้อยกว่า และต้องยอมรับความจริงอย่างหนึ่งว่า ตัวแทนครัวเรือนที่ไม่ตอบเรื่องรายได้ในแต่ละชุดนั้น อาจมีบางหน่วยที่ดูแล้วค่ารายได้ อาจจะผิดปกติ แต่การศึกษาครั้งนี้เป็นการศึกษาถึงวิธีการประมาณค่าเท่านั้น จึงไม่ได้หาค่าที่คิดว่าผิดปกติออกไป ค่าเฉลี่ยของร้อยละของผลต่างจากค่าจริงอาจจะมีค่ามากในที่นี้ ได้ทำการเปรียบเทียบค่าเฉลี่ยของร้อยละของผลต่างระหว่างค่าประมาณกับค่าจริงของรายได้จากตัวแทนหน่วยที่ไม่ตอบเรื่องรายได้ 2 ชุดคือ

1. ตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ที่เลือกขึ้นมาโดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัวเรือนจำนวน 61 ครัวเรือนจากครัวเรือนที่ให้คำตอบเรื่องรายได้ 1,436 ครัวเรือนในแต่ละหน่วยได้นำค่าประมาณหักออกจากรายได้จริงแล้วหารด้วยค่าประมาณคูณด้วยร้อยละผลที่ได้ไม่คิดเครื่องหมายแล้วจึงนำร้อยละของผลต่างทั้ง 61 ครัวเรือน มาหาค่าเฉลี่ยผลลัพธ์สำหรับค่าเฉลี่ยจากชั้นภูมิค่าสุ่มจากชั้นภูมิและจากสมการถดถอยได้แสดงไว้ในสภมภ์ที่ 6 7 และ 8 ตามลำดับของตารางที่ 18

2. ตัวแทนหน่วยที่ไม่ตอบในเรื่องรายได้ที่เลือกขึ้นมาโดยวิธีสุ่ม จำนวน 74 ครัวเรือนจากครัวเรือนที่ให้คำตอบเรื่องรายได้ 1,436 ครัวเรือน ในการหาค่าเฉลี่ยของร้อยละของผลต่างจากรายได้จริงของรายได้ทำเช่นเดียวกับข้อ 1 ผลลัพธ์สำหรับค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิและจากสมการถดถอยได้แสดงไว้ในสภมภ์ที่ 6 7 และ 8 ตามลำดับของตารางที่ 19

ตารางที่ 18 แสดงค่าจริงกับค่าประมาณรายได้ที่ประมาณจากวิธีต่าง ๆ และร้อยละของผลต่างจากรายได้จริง ของตัวแทนหน่วยที่ไม่ตอบเรื่องรายได้ ซึ่งเลือกขึ้นมาโดยอาศัยปัจจัยของหัวหน้าครัวเรือน

ลำดับที่	ค่าจริง	ค่าที่คำนวณได้จาก			ร้อยละของผลต่างจากรายได้จริง		
		ค่าเฉลี่ยจากชั้นภูมิ	ค่าสุ่มจากชั้นภูมิ	สมการถดถอย	ค่าเฉลี่ยจากชั้นภูมิ	ค่าสุ่มจากชั้นภูมิ	สมการถดถอย
1	10,000	9,686	12,000	4,611	3.14	20.00	53.89
2	10,000	9,686	25,000	5,379	3.14	150.00	46.21
3	2,500	2,396	3,350	1,576	4.16	34.00	36.96
4	13,000	9,686	8,000	5,640	25.49	38.46	56.62
5	10,000	14,945	10,000	9,431	49.45	0.00	5.68
6	25,000	9,686	4,000	5,307	61.25	84.00	78.77
7	15,000	14,945	12,000	10,872	0.36	20.00	27.52
8	7,100	9,686	8,500	6,216	36.42	19.72	12.45
9	10,000	4,069	8,000	2,723	59.31	20.00	72.77
10	7,500	4,069	2,150	2,976	45.74	71.33	60.32
11	4,000	4,069	2,500	2,789	1.72	37.50	30.29
12	1,800	4,069	2,080	2,115	126.05	15.55	17.53
13	2,000	2,396	3,000	2,663	19.80	50.00	33.17

## ตารางที่ 18 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่คำนวณได้จาก			ร้อยละของผลต่างจากค่าจริง		
		ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย	ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย
14	7,500	4,069	3,540	3,623	45.74	52.80	51.69
15	3,000	4,069	2,000	3,161	35.63	33.33	5.37
16	3,000	2,396	1,560	1,996	20.13	48.00	33.45
17	3,000	4,069	18,000	3,129	35.63	500.00	4.30
18	4,000	4,069	3,500	2,816	1.72	12.50	29.61
19	1,000	4,069	4,500	2,683	306.90	350.00	168.31
20	5,000	4,069	1,500	2,092	18.62	70.00	58.16
21	580	4,069	3,580	1,444	601.55	517.24	148.90
22	2,000	9,686	6,000	5,492	245.92	114.28	96.15
23	7,000	2,396	3,400	2,207	65.77	51.42	71.04
24	25,000	9,686	10,000	5,898	61.25	60.00	76.41
25	3,000	4,069	5,500	3,243	35.63	83.33	8.11
26	6,000	4,069	4,830	3,864	32.18	19.50	35.60
27	6,000	9,686	15,000	5,092	61.43	150.00	15.14
28	20,000	4,069	3,530	3,660	79.65	82.35	81.70
29	3,000	4,069	3,800	3,308	35.63	26.66	10.26
30	3,500	4,069	15,000	2,141	16.26	328.57	38.83
31	1,000	4,069	1,200	2,190	306.90	20.00	119.00
32	2,500	4,069	2,400	4,184	62.76	4.00	67.37
33	4,000	4,069	4,500	3,826	1.72	12.50	4.35
34	25,000	9,686	5,000	4,598	61.25	80.00	81.60
35	3,000	4,069	3,800	3,754	35.63	26.66	25.12
36	1,500	4,069	1,000	2,573	171.26	33.33	71.52
37	2,500	4,069	2,000	3,488	62.76	20.00	39.52

ตารางที่ 18 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่คำนวณได้จาก			ร้อยละของผลต่างจากค่าจริง		
		ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย	ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย
38	1,800	4,069	4,700	3,314	126.05	161.11	84.12
39	9,000	4,069	800	4,320	54.78	91.11	52.00
40	3,000	4,069	2,000	3,440	35.63	33.33	14.66
41	4,000	4,069	3,500	1,619	1.72	12.50	59.52
42	3,200	4,069	2,000	2,337	27.15	37.50	26.98
43	20,000	4,069	6,000	1,735	79.65	70.00	91.33
44	9,000	4,069	11,000	4,218	54.78	22.22	53.13
45	13,500	9,686	3,500	5,720	28.25	74.09	57.63
46	30,000	9,686	3,100	9,944	67.71	89.66	66.85
47	2,910	9,686	5,000	2,363	232.85	71.82	18.81
48	2,400	4,069	4,440	2,516	69.54	85.00	4.83
49	4,000	4,069	1,500	3,697	1.72	62.50	7.59
50	9,000	4,069	3,500	5,063	54.78	61.11	43.75
51	5,000	2,396	3,000	3,498	52.08	40.00	30.03
52	3,000	4,069	3,000	3,757	35.63	0.00	25.24
53	1,500	4,069	3,000	2,076	171.26	100.00	38.39
54	2,500	4,069	2,700	1,939	62.76	8.00	22.45
55	1,500	2,396	3,000	2,381	59.73	100.00	58.74
56	750	4,000	3,000	1,820	442.53	300.00	142.67
57	500	2,396	2,000	1,193	379.20	300.00	138.58
58	6,500	9,686	4,000	3,694	49.02	38.46	43.17
59	1,560	2,396	3,500	1,389	53.59	124.36	10.93
60	3,000	2,396	600	3,179	20.13	80.00	5.95
61	80	2,396	3,000	1,645	2,895.00	3,650.00	1,956.60
ค่าเฉลี่ย					129.40	145.40	80.78



ตารางที่ 19 แสดงค่าจริงกับค่าประมาณรายได้ที่ประมาณจากวิธีต่าง ๆ และร้อยละของผลต่าง  
จากรายได้จริงของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ซึ่งเลือกขึ้นมาโดยสุ่ม

ลำดับที่	ค่าจริง	ค่าที่คำนวณได้จาก			ร้อยละของผลต่างจากค่าจริง		
		ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย	ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย
1	3,000	4,099	3,500	2,261	36.63	16.66	24.63
2	3,500	4,099	4,000	3,196	17.11	14.28	8.68
3	2,500	4,099	2,000	4,424	63.96	20.00	76.96
4	9,000	2,398	1,800	2,718	73.36	80.00	69.80
5	34,500	9,928	10,000	7,735	71.22	71.01	77.58
6	300	4,099	2,000	1,645	1,266.33	556.66	448.33
7	500	2,398	3,000	1,637	379.60	500.00	227.40
8	2,000	4,099	580	2,371	104.95	71.00	18.55
9	4,000	4,099	4,600	3,915	2.47	15.00	2.12
10	3,000	9,928	10,000	3,119	230.93	233.33	3.96
11	35,000	14,204	45,000	11,621	59.42	28.57	66.80
12	31,000	9,928	6,500	7,993	67.97	79.03	74.22
13	1,500	2,398	2,010	939	59.86	34.00	37.40
14	2,110	2,398	4,600	1,757	13.65	118.01	16.73
15	3,000	4,099	600	2,980	36.63	80.00	0.66
16	800	4,099	3,500	3,906	412.38	337.50	388.25
17	900	4,099	3,000	2,175	355.44	233.33	141.66
18	3,540	4,099	9,000	3,691	15.79	154.24	4.26
19	12,000	4,099	5,000	4,092	65.84	58.33	65.90
20	2,000	1,909	3,000	1,176	4.55	50.00	41.20
21	3,500	4,099	5,000	2,727	17.11	42.86	22.08
22	2,000	4,099	3,000	3,636	104.95	50.00	81.80
23	4,900	4,099	1,890	3,780	16.34	61.43	22.85

ตารางที่ 19 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่คำนวณได้จาก			ร้อยละของผลต่างจากค่าจริง		
		ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย	ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย
24	2,600	2,398	450	1,531	7.77	82.69	41.11
25	13,000	9,928	25,000	5,763	23.63	92.31	55.67
26	5,000	4,099	7,610	2,565	18.02	52.20	48.7
27	1,350	9,928	9,000	6,388	635.41	566.67	373.18
28	2,000	4,099	5,000	3,421	104.95	150.00	71.05
29	2,500	4,099	1,500	3,565	63.96	40.00	42.60
30	2,000	2,398	1,500	1,598	19.90	25.00	20.10
31	6,200	9,928	20,000	9,556	60.13	222.58	54.13
32	4,200	9,928	3,100	5,088	136.38	26.19	21.14
33	2,000	4,099	2,100	2,617	104.95	5.00	30.85
34	1,000	4,099	3,000	3,399	309.90	200.00	239.90
35	4,600	9,928	10,000	8,243	115.82	117.39	79.19
36	15,000	4,099	3,000	3,443	72.67	80.00	77.04
37	2,000	2,398	1,500	1,600	19.90	25.00	20.00
38	1,500	4,099	2,200	2,601	173.26	46.67	73.4
39	20,000	4,099	7,000	3,653	79.50	65.00	81.73
40	4,300	9,928	5,000	5,723	130.88	16.28	33.09
41	6,000	4,099	2,000	4,928	31.68	66.67	17.86
42	8,000	4,099	5,000	3,589	48.76	37.50	55.14
43	3,000	4,099	4,000	2,780	36.63	33.33	7.33
44	1,000	2,398	1,500	1,768	139.80	50.00	76.80
45	1,000	2,398	2,000	2,156	139.80	100.00	115.60
46	3,000	4,099	1,660	3,243	36.63	44.67	8.10
47	2,500	4,099	6,500	2,924	63.96	160.00	16.96
48	2,000	2,398	3,000	1,536	19.90	50.00	23.20
49	4,000	4,099	4,000	3,296	2.48	0.00	17.60
50	5,000	9,928	6,000	5,858	98.56	20.00	17.16

ตารางที่ 19 (ต่อ)

ลำดับที่	ค่าจริง	ค่าที่คำนวณได้จาก			ร้อยละของผลต่างจากค่าจริง		
		ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย	ค่าเฉลี่ยจาก ชั้นภูมิ	ค่าสุ่มจาก ชั้นภูมิ	สมการ ถดถอย
51	2,100	4,099	3,000	3,964	95.19	42.86	88.76
52	1,260	2,398	3,000	1,482	90.32	138.10	17.62
53	3,250	9,928	3,000	5,670	205.48	7.69	74.46
54	6,000	4,099	3,000	3,333	31.68	50.00	44.45
55	1,000	14,204	12,000	12,670	1,320.04	1,100.00	1,167.00
56	2,000	4,099	1,900	3,361	104.95	5.00	68.05
57	3,300	4,099	3,000	2,837	24.21	9.09	14.03
58	1,600	4,099	1,500	1,417	156.18	6.25	11.43
59	6,500	9,928	4,000	5,484	52.74	38.46	15.63
60	4,000	9,928	5,000	7,210	148.20	25.00	80.25
61	6,000	4,099	2,000	4,647	31.68	66.67	22.55
62	7,000	4,099	7,000	4,595	41.44	0.00	34.35
63	120	9,928	3,000	2,677	8,173.33	2,400.00	2,130.83
64	5,200	4,099	7,500	4,266	21.17	44.23	17.96
65	1,000	4,099	580	2,819	309.90	42.00	181.90
66	2,800	4,099	550	3,722	46.39	80.36	32.93
67	5,000	2,398	1,800	3,416	52.04	64.00	31.68
68	3,000	4,099	4,000	3,194	36.63	33.33	6.46
69	2,540	4,099	4,000	2,790	61.38	57.48	9.84
70	830	4,099	2,500	1,957	393.86	201.20	135.78
71	2,000	4,099	4,000	3,652	104.95	100.00	82.60
72	2,000	4,099	6,500	3,186	104.95	225.00	59.30
73	4,500	4,099	3,000	3,388	8.91	33.33	24.71
74	3,000	2,398	3,000	3,155	20.06	0.00	5.16
ค่าเฉลี่ย					240.64	135.95	108.08

จากตารางที่ 18 ที่แสดงร้อยละของผลต่างจากค่าจริงของรายได้ของตัวแทนจำนวน 61 ครว เรือนที่เลือกขึ้นมาโดยอาศัยปัจจัยพื้นฐานของหัวหน้าครว เรือน ในแต่ละวิธีคือ วิธีประมาณจาก ค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิ และจากสมการถดถอย ผลปรากฏว่าวิธีประมาณค่ารายได้ที่สมมติ ว่าขาดหายไปโดยวิธีสมการถดถอย ให้ค่าเฉลี่ยร้อยละของผลต่างจากค่าจริงน้อยที่สุดคือร้อยละ 80.78 รองลงมาได้แก่ วิธีประมาณโดยใช้ค่าเฉลี่ยจากชั้นภูมิ คิดเป็นร้อยละ 129.40 และมีผลต่างจากค่าจริงมากที่สุดคือ วิธีประมาณโดยใช้ค่าสุ่มจากชั้นภูมิคิดเป็นร้อยละ 145.50

จากตารางที่ 19 ที่แสดงร้อยละของผลต่างจากค่าจริงของรายได้ของตัวแทนจำนวน 71 ครว เรือนที่เลือกขึ้นมาโดยสุ่ม ในแต่ละวิธีเหมือนข้างต้น ผลปรากฏว่า วิธีประมาณค่ารายได้ที่สมมติ ว่าขาดหายไป วิธีสมการถดถอยให้ค่าเฉลี่ยร้อยละของผลต่างจากค่าจริงน้อยที่สุดคือคิด เป็นร้อยละ 108.08 รองลงมาได้แก่ วิธีค่าสุ่มจากชั้นภูมิ คิดเป็นร้อยละ 135.95 และวิธีที่ทำให้มีค่าเฉลี่ยของ ร้อยละของผลต่างจากค่าจริงมากที่สุดคือ วิธีประมาณ โดยใช้ค่าเฉลี่ยจากชั้นภูมิคือคิด เป็นร้อยละ 240.64

### 3.5 การเปรียบเทียบค่าจริงและค่าประมาณของรายได้และเปรียบเทียบวิธีการประมาณทั้ง 3 วิธี

#### 1. การเปรียบเทียบค่าจริงและค่าประมาณของรายได้

นำค่าประมาณของรายได้ที่สมมติว่าขาดหายไป ที่ประมาณโดยใช้ค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิ และจากสมการถดถอย มาทดสอบความแตกต่างกับค่าจริงของรายได้แบบจับคู่สิ่ง ทดลอง เพื่อดูว่า ค่าประมาณที่คำนวณได้มีความคลาดเคลื่อนจากค่าจริงของรายได้ อย่างมีนัยสำคัญ ทางสถิติหรือไม่ ในที่นี้ได้แบ่งการทดสอบออกเป็น 2 ชุด คือ

ก. ทดสอบกับขนาดตัวอย่างย่อยของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ซึ่ง เลือกขึ้นมา โดยอาศัยปัจจัยพื้นฐานของหัวหน้าครว เรือนจำนวน 61 ครว เรือน ในการทดสอบความแตกต่างได้นำ ผลการประมาณค่าจากตารางที่ 18 ซึ่งแสดงผลการประมาณจาก ค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิ และจาก สมการถดถอย ในสมมติที่ 3, 4 และ 5 ตามลำดับ มาทดสอบแบบจับคู่สิ่งทดลองกับค่า จริงของรายได้ ซึ่งอยู่ในสมมติที่ 2 ค่า t-statistics และผลการทดสอบ แสดงไว้ในตารางที่ 20

ข. ทดสอบกับขนาดตัวอย่างย่อยของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ ซึ่ง เลือกขึ้นมา โดยวิธีสุ่มจำนวน 74 ครว เรือน ในกรทดสอบความแตกต่างได้นำผลการประมาณค่าจากตารางที่ 19 ซึ่งแสดงผลการประมาณ จาก ค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิ และจาก สมการถดถอย ในสมมติที่ 3, 4 และ 5 ตามลำดับ มาทดสอบแบบจับคู่สิ่งทดลอง กับค่าจริงของรายได้ ซึ่งอยู่ใน สมมติที่ 2 ค่า t-statistics และผลการทดสอบ แสดงไว้ ในตารางที่ 21



2. การเปรียบเทียบวิธีที่ใช้ในการประมาณรายได้ที่ขาดหายไป ทั้ง 3 วิธี

นำค่าประมาณที่ได้จากทั้ง 3 วิธี คือวิธีใช้ค่าเฉลี่ยจากชั้นภูมิ เมื่อเลือกตัวอย่างแล้ว ใช้ค่าสุ่มจากชั้นภูมิ เมื่อเลือกตัวอย่างแล้ว และวิธีวิเคราะห์ความถดถอยมาทดสอบดูว่าค่าประมาณที่ได้จากแต่ละวิธีมีความแตกต่างกันหรือไม่ ในที่นี้ได้ทำการทดสอบกับข้อมูล 2 ชุดคือ

ก. ทดสอบ กับขนาดตัวอย่างย่อยของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ ซึ่ง เลือกขึ้นมาโดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัวเรือนจำนวน 61 ครัวเรือนในการทดสอบความแตกต่างของทั้ง 3 วิธีได้นำผลการประมาณค่าจากตารางที่ 18 ซึ่งแสดงผลการประมาณจากการใช้วิธีหาค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิ และวิธีวิเคราะห์ความถดถอยในสัปดาห์ที่ 3 4 และ 5 ตามลำดับ มาทดสอบแบบจับคู่สิ่งทดลอง คือ

1) ทดสอบดูว่า วิธีประมาณค่ารายได้ จากการใช้ค่าเฉลี่ยจากชั้นภูมิและค่าสุ่มจากชั้นภูมิ มีความแตกต่างกันหรือไม่

2) ทดสอบดูว่า วิธีประมาณค่ารายได้จากการใช้ ค่าเฉลี่ยจากชั้นภูมิและจากการวิเคราะห์ความถดถอย มีความแตกต่างกันหรือไม่

3) ทดสอบดูว่า วิธีประมาณค่ารายได้ จากการใช้ค่าสุ่มจากชั้นภูมิและจากการวิเคราะห์ความถดถอย มีความแตกต่างกันหรือไม่

ค่า  $t$ -statistics ที่คำนวณได้และผลการทดสอบ แสดงไว้ในตารางที่ 20

ข. ทดสอบกับขนาดตัวอย่างย่อยของตัวแทนหน่วยที่ไม่ตอบ เรื่องรายได้ซึ่งเลือกขึ้นมาโดยวิธีสุ่ม จำนวน 74 ครัวเรือน นำผลการประมาณค่ารายได้จากตารางที่ 19 ซึ่งแสดงผลการประมาณจากการใช้ค่าเฉลี่ยจากชั้นภูมิ ค่าสุ่มจากชั้นภูมิและวิธีวิเคราะห์ความถดถอยในสัปดาห์ที่ 3 4 และ 5 ตามลำดับ มาทดสอบความแตกต่างของทั้ง 3 วิธี เช่นเดียวกับที่ได้อธิบายไว้ในข้อ ก.

ค่า  $t$ -statistics ที่คำนวณได้ และผลการทดสอบแสดงไว้ในตารางที่ 21

ตารางที่ 20 แสดงค่าสถิติ t ที่คำนวณได้จากการทดสอบสมมติฐานเกี่ยวกับความแตกต่างระหว่างรายได้ที่ประมาณได้จาก 3 วิธี กับค่าจริงของรายได้และระหว่างรายได้ที่ประมาณได้จากแต่ละวิธีของตัวแทนของครัวเรือนที่ไม่ตอบเรื่องรายได้โดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัวเรือน

รายได้	ค่า t-statistics		
	ค่าประมาณจาก ค่าเฉลี่ยจากชั้นภูมิ	ค่าประมาณจาก ค่าสุ่มจากชั้นภูมิ	ค่าประมาณจาก สมการถดถอย
ค่าจริง	1.613	1.590	3.901**
ค่าประมาณจาก ค่าเฉลี่ยจากชั้นภูมิ		0.591	7.271**
ค่าประมาณจาก ค่าสุ่มจากชั้นภูมิ			2.679**

หมายเหตุ : \* หมายถึง ปฏิเสธ  $H_0$  ที่ระดับนัยสำคัญ .05

\*\* หมายถึง ปฏิเสธ  $H_0$  ที่ระดับนัยสำคัญ .01

d.f = 60

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 21 แสดงค่าสถิติ  $t$  ที่คำนวณได้จากการทดสอบสมมติฐานเกี่ยวกับความแตกต่างระหว่างรายได้ที่ประมาณได้จาก 3 วิธีกับค่าจริงของรายได้ และระหว่างรายได้ที่ประมาณได้จากแต่ละวิธีของตัวแทนครัวเรือนที่ไม่ตอบเรื่องรายได้ที่เลือกขึ้นมาโดยวิธีสุ่ม

รายได้	ค่า $t$ -statistics		
	ค่าประมาณจาก ค่าเฉลี่ยจากชั้นภูมิ	ค่าประมาณจาก ค่าสุ่มจากชั้นภูมิ	ค่าประมาณจาก สมการถดถอย
ค่าจริง	-0.2934	0.005	1.648
ค่าประมาณจาก ค่าเฉลี่ยจากชั้นภูมิ		0.376	7.306**
ค่าประมาณจาก ค่าสุ่มจากชั้นภูมิ			1.968*

หมายเหตุ : \* หมายถึง ปฏิเสธ  $H_0$  ที่ระดับนัยสำคัญ .05

\*\* หมายถึง ปฏิเสธ  $H_0$  ที่ระดับนัยสำคัญ .01

degree of freedom = 73

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

จากตารางที่ 20 และตารางที่ 21 เมื่อทดสอบความแตกต่างระหว่างค่าจริงกับค่าประมาณ รายได้ที่ประมาณโดยใช้ค่าเฉลี่ยจากชั้นภูมิ ค่าลุ่มจากชั้นภูมิ และวิธีวิเคราะห์ความถดถอย และการทดสอบความแตกต่างระหว่างวิธีที่ใช้ในการประมาณรายได้ สรุปได้ดังนี้

1. สำหรับตัวแทนหน่วยที่ไม่ตอบเรื่องรายได้ ซึ่งเลือกมาจากหน่วยที่ให้คำตอบเรื่อง รายได้จำนวนทั้งสิ้น 1,436 ครว เรือน โดยอาศัยปัจจัยพื้นฐานของหัวหน้าครัว เรือนแล้วได้ 61 ครัว เรือน

ก. ค่าประมาณรายได้ที่สมมติว่าขาดหายไปทั้ง 61 ครัว เรือน โดยวิธีใช้ค่าเฉลี่ย จากชั้นภูมิ ค่าลุ่มจากชั้นภูมิ และวิธีที่ใช้การวิเคราะห์ความถดถอย เมื่อนำมาทดสอบกับค่าจริงของ รายได้ที่สมมติว่าขาดหายไปปรากฏว่าค่าประมาณที่ได้จากวิธีที่ใช้ค่าเฉลี่ยจากชั้นภูมิ และวิธีที่ใช้ค่า ลุ่มจากชั้นภูมิไม่แตกต่างจากค่าจริงของรายได้ ส่วนการใช้วิธีวิเคราะห์ความถดถอยให้ค่าประมาณ ของรายได้ แตกต่างจากค่าจริง ที่ระดับนัยสำคัญ .05

ข. ค่าประมาณรายได้ที่ประมาณโดยวิธีที่ใช้ค่าเฉลี่ยจากชั้นภูมิและวิธีที่ใช้ค่าลุ่มจาก ชั้นภูมิ ให้ค่าประมาณไม่แตกต่างกันอย่างมีนัยสำคัญ ส่วนค่าประมาณรายได้ที่ประมาณโดยวิธีที่ใช้ค่า เฉลี่ยจากชั้นภูมิกับวิธีที่ใช้การวิเคราะห์ความถดถอยและวิธีที่ใช้ค่าลุ่มจากชั้นภูมิกับวิธีที่ใช้การวิเคราะห์ ความถดถอยให้ค่าประมาณแตกต่างกันที่ระดับนัยสำคัญ .05

2. สำหรับตัวแทนหน่วยที่ไม่ตอบเรื่องรายได้ ซึ่งเลือกมาจากหน่วยที่ให้คำตอบเรื่องรายได้ จำนวนทั้งสิ้น 1,436 ครว เรือน โดยวิธีลุ่มแล้วได้ 74 ครัว เรือน

ก. ค่าประมาณรายได้ที่ประมาณโดยวิธีใช้ค่าเฉลี่ยจากชั้นภูมิ ค่าลุ่มจากชั้นภูมิและวิธี ที่ใช้การวิเคราะห์ความถดถอย เมื่อนำมาทดสอบกับค่าจริง ปรากฏว่าค่าประมาณที่ได้จากทั้ง 3 วิธี ไม่แตกต่างจากค่าจริงของรายได้ที่สมมติว่าขาดหายไปที่ระดับนัยสำคัญ .05

ข. ค่าประมาณรายได้ที่ประมาณโดยวิธีที่ใช้ค่าเฉลี่ยจากชั้นภูมิและวิธีที่ใช้ค่าลุ่มจาก ชั้นภูมิ ให้ค่าประมาณไม่แตกต่างกันอย่างมีนัยสำคัญ ส่วนค่าประมาณรายได้ที่ประมาณโดยวิธีที่ใช้ค่า เฉลี่ยจากชั้นภูมิกับวิธีที่ใช้การวิเคราะห์ความถดถอยและวิธีใช้ค่าลุ่มจากชั้นภูมิกับวิธีที่ใช้การวิเคราะห์ ความถดถอยให้ค่าประมาณแตกต่างกันที่ระดับนัยสำคัญ 0.05