

### วิธีทดสอบฐานคติเบื้องต้น

(Basic assumption testing)

ประเด็นในการทดสอบฐานคติเบื้องต้นดังกล่าวไว้ในบทที่ 5 นั้นมีด้วยกัน 5 ประเด็น ดังจะเสนอรายละเอียดวิธีการทดสอบดังนี้คือ

1. ข้อมูลต้องอยู่ในมาตราวัดอัตราภาคหรือมาตราวัดอัตราส่วน (interval or ratio data)

โดยหลักการ ตัวแปรทุกตัวในแบบจำลองสาเหตุและผลจะต้องมีการวัดค่าในมาตราวัดอันตรภาค (interval scale) หรือมาตราวัดอัตราส่วน (ratio scale) หรือมาตราวัดอื่นที่พอเทียบเคียงกับสองระบบดังกล่าว ฐานคตินี้หากมองโดยผิวเผินก็ไม่น่าจะก่อให้เกิดปัญหาแต่ประการใด แต่ถ้าพิจารณาให้รอบคอบแล้วก็จะพบว่า ฐานคตินี้ก่อให้เกิดปัญหาต่อการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์อยู่ไม่น้อย ทั้งนี้เพราะตัวแปรส่วนใหญ่ที่ใช้ในศาสตร์ทั้งสองแขนงนี้มักมีมาตราวัดอยู่เพียงไม่เกินขั้นจัดลำดับ (ordinal scale) เท่านั้น (ประชัย เปี่ยมสมบูรณ์, 2524 : 62) ซึ่งถ้าเป็นเช่นนั้นแล้วก็จะทำให้การวิเคราะห์ข้อมูลอาจคลาดเคลื่อนได้ ดังนั้น ผู้วิจัยก็ได้สร้างมาตราวัดสำหรับตัวแปรที่จะใช้สถิติถดถอย ซึ่งตัวแปรเหล่านี้เราอาจเรียกว่าเป็นตัวแปรต่อเนื่อง (continuous) ข้อมูลสำหรับการวิจัยที่เก็บรวบรวมมาก็ถือว่าผ่านฐานคติข้อนี้

2. จะต้องไม่มีการระบุผิด (no specification error)

2.1 ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามต้องมีลักษณะเป็นเส้นตรง

2.2 จะต้องไม่มีตัวแปรอิสระใด ๆ ที่มีความสำคัญต่อการอธิบายตัวแปรตามอยู่นอกตัวแบบที่ใช้ในการศึกษา

2.3 จะต้องไม่มีตัวแปรอิสระใดๆ ที่ไม่มีความสำคัญต่อการอธิบายตัวแปรตามอยู่ในตัวแบบที่ใช้ในการศึกษา

2.4 ค่าความแปรปรวนผิดพลาดต้องมีค่าคงที่ (homoscedasticity :  $E(e^2 | i) = \sigma^2$ )

## 2.1 ความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามจะต้องมีลักษณะเป็นเส้นตรง (linearity)

การทดสอบความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามว่า มีลักษณะเป็นเส้นตรงหรือไม่นั้น ก็เป็นฐานคติหนึ่งสำหรับการวิเคราะห์ถดถอย การทดสอบก็เพื่อให้มั่นใจว่าความสัมพันธ์ระหว่างตัวแปรไม่ใช่เส้นโค้ง (curvilinearity) นั้นมีความจำเป็นมากในการวิเคราะห์ตัวแปรซึ่งมีค่าต่อเนื่อง (continuous variable) เพราะถ้าใช้สมการเส้นตรงกับข้อมูลที่ เป็นเส้นโค้ง (curvilinearity data) ผลลัพธ์ คือ ค่าสัมประสิทธิ์บ่งบอก (coefficient of determination หรือ  $R^2$ ) ย่อมมีขนาดต่ำกว่าความเป็นจริง ซึ่งเป็น การบิดเบือนความเป็นจริงของปรากฏการณ์ที่กำลังศึกษา ด้วยเหตุนี้ นักวิจัยจึงต้องทำการทดสอบก่อนว่า ความเบี่ยงเบนจากเส้นตรง (Deviation from linearity) มีค่านัยสำคัญทางสถิติหรือไม่ (วิล วังศ์สืบชาติ, 2528 : 59)

ตามปกติแล้วความสัมพันธ์ระหว่างตัวแปรต่าง ๆ ทางด้านสังคมศาสตร์มักจะได้รับการ assume ว่ามีลักษณะที่เป็นเส้นตรงแน่นอนที่สุด ฐานคติเบื้องต้นในลักษณะนี้อาจจะไม่เป็นการถูกต้องเสมอไป แต่อย่างน้อยที่สุดการยอมรับข้อตกลงนี้ก็ได้รับการยืนยันอย่างแน่ชัดว่ามีความเชื่อถือได้มากก็เพราะว่า ประการแรก มีความสัมพันธ์ระหว่างตัวแปรอย่างมากมายที่พบจากข้อมูลว่าเป็นเส้นตรง ประการที่สอง ความสัมพันธ์ที่เป็นเส้นตรงจะเข้ากับหลักการวิทยาศาสตร์ในประเด็นของการประหยัดอย่างมาก (parsimonious) (Lewis-beck, 1980 : 13) อีกทั้งการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามในการวิเคราะห์ถดถอยได้ยึดหลักกำลังสองที่ต่ำสุดระหว่างตัวแปรอิสระและตัวแปรตาม จึงต้องมีลักษณะเป็นเส้นตรงอย่างหลีกเลี่ยงไม่ได้ สำหรับการตรวจสอบว่าความสัมพันธ์ระหว่างตัวแปรจะเป็นเส้นตรงหรือไม่นั้น จะใช้วิธีการทดสอบโดยการวิเคราะห์ความแปรปรวนของตัวแปรที่ละคู่ ระหว่างตัวแปรอิสระและตัวแปรตาม ซึ่งผลของการคำนวณจะให้ค่านัยสำคัญของ F-test จากค่าความเป็นเส้นตรง (linearity) และค่านัยสำคัญของ F-test จากค่าความเบี่ยงเบนออกจากเส้นตรง (deviation from linearity) (ประสิทธิ์ สันติกาญจน์, 2530 : 132)

สำหรับในการนิยามนั้น จะมีการตั้งสมมติฐานและเปรียบเทียบค่าที่จุดวิกฤต (critical values) ซึ่งมีค่านัยสำคัญที่ตั้งขึ้น ( $\alpha$ ) มีค่าเท่ากับ 0.01 ในการทดสอบ หากว่าค่านัยสำคัญที่คำนวณได้ (P) มีค่านัยสำคัญน้อยกว่าที่ตั้งขึ้น ( $\alpha$ ) แสดงว่า significant

แต่หากว่าค่าที่คำนวณได้มากกว่าค่าที่ได้ตั้งขึ้นแสดงว่า not significant ดังแสดงไว้ในตารางการวิเคราะห์ความแปรปรวน (analysis of variance) ดังต่อไปนี้

ตารางที่ 20

แสดงผลการวิเคราะห์ความแปรปรวนระหว่างโอกาสความก้าวหน้าในการทำงาน  
กับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	266.858	1	266.858	6.932	0.0090

with fewer than three groups, The relationship is linear

จากการวิเคราะห์ความแปรปรวนระหว่างโอกาสความก้าวหน้าในการทำงาน  
กับความตั้งใจที่จะออกจากองค์การ พบว่า มีความสัมพันธ์เชิงเส้นตรง เพราะว่ากลุ่มในการ  
วิเคราะห์น้อยกว่า 3 กลุ่ม จึงถือว่ามีความสัมพันธ์เชิงเส้นตรงโดยอัตโนมัติ

ตารางที่ 21

แสดงผลการวิเคราะห์ความแปรปรวนระหว่างความเป็นอิสระในการทำงาน  
กับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	811.551	1	811.554	22.469	0.0090

with fewer than three groups, the relationship is linear

จากการวิเคราะห์ความแปรปรวนระหว่างความเป็นอิสระในการทำงาน กับความตั้งใจที่จะออกจากองค์การ พบว่า มีความสัมพันธ์เชิงเส้นตรง เพราะว่ากลุ่มในการวิเคราะห์น้อยกว่า 3 กลุ่ม จึงถือว่ามีความสัมพันธ์เชิงเส้นตรงโดยอัตโนมัติ

### ตารางที่ 22

แสดงผลของการวิเคราะห์ความแปรปรวนระหว่างการสื่อสารในองค์การ  
กับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	348.445	2	174.222	4.548	0.0116
Linearity	318.608	1	318.608	8.317	0.0043
DEV. from linearity	29.836	1	29.836	0.779	0.3784

A) ทดสอบว่าความสัมพันธ์ระหว่างการสื่อสารในองค์การกับความตั้งใจที่จะออกจากองค์การมีลักษณะเป็นเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเป็นเส้นตรง

$$P = 0.0043 < X = 0.01$$

แสดงว่ามีนัยสำคัญจึงยอมรับ  $H_1$  ที่ว่าลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นเส้นตรง

B) ทดสอบว่าความสัมพันธ์ระหว่างการสื่อสารในองค์การและความตั้งใจที่จะออกจากองค์การเป็นเส้นตรงหรือไม่

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเบี่ยงเบนออกจากความเป็นเส้นตรง

$$P = 0.3784 < X = 0.01$$

แสดงว่าไม่มีนัยสำคัญจึงยอมรับ  $H_0$  ที่ว่า ลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

### ตารางที่ 23

แสดงผลของการวิเคราะห์ความแปรปรวนระหว่างการรวมอำนาจ  
กับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	893.794	2	446.897	12.443	0.0000
Linearity	790.879	1	790.879	22.020	0.0000
DEV. from linearity	102.915	1	102.915	2.865	0.0919

A) ทดสอบว่าความสัมพันธ์ระหว่างการรวมอำนาจและความตั้งใจที่จะออกจากองค์การมีลักษณะเป็นเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเป็นเส้นตรง

$$P = 0.0000 < X = 0.01$$

แสดงว่ามีนัยสำคัญจึงยอมรับ  $H_1$  ที่ว่าลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นเส้นตรง

B) ทดสอบว่าความสัมพันธ์ระหว่างการรวมอำนาจและความตั้งใจที่จะออกจากองค์การมีการเบี่ยงเบนออกจากเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเบี่ยงเบนออกจากความเป็นเส้นตรง

$$P = 0.0919 < X = 0.01$$

แสดงว่าไม่มีนัยสำคัญจึงยอมรับ  $H_0$  ที่ว่า ลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

#### ตารางที่ 24

แสดงผลของการวิเคราะห์ความแปรปรวนระหว่างรายได้  
กับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	1008.729	6	168.122	4.664	0.0002
Linearity	635.259	1	635.259	17.624	0.0000
DEV. from linearity	373.471	5	74.694	2.072	0.0699

A) ทดสอบว่าความสัมพันธ์ระหว่างรายได้และความตั้งใจที่จะออกจากองค์การ มีลักษณะเป็นเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเป็นเส้นตรง

$$P = 0.0000 < X = 0.01$$

แสดงว่ามีนัยสำคัญจึงยอมรับ  $H_1$  ที่ว่าลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นเส้นตรง

B) ทดสอบว่าความสัมพันธ์ระหว่างรายได้และความตั้งใจที่จะออกจากองค์การ มีการเบี่ยงเบนออกจากเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสอง ไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเบี่ยงเบนออกจากความเป็นเส้นตรง

$$P = 0.0699 < X = 0.01$$

แสดงว่าไม่มีนัยสำคัญจึงยอมรับ  $H_0$  ที่ว่า ลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสอง ไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

#### ตารางที่ 25

แสดงผลของการวิเคราะห์ความแปรปรวนระหว่างสวัสดิการในการทำงาน  
กับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	1297.271	2	648.635	18.995	0.0000
Linearity	1281.203	1	1281.203	37.520	0.0000
DEV. from linearity	16.068	1	16.068	0.471	0.4934

A) ทดสอบว่าความสัมพันธ์ระหว่างสวัสดิการในการทำงานและความตั้งใจที่จะออกจากองค์การมีลักษณะเป็นเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสอง ไม่ได้เป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเป็นเส้นตรง

$$P = 0.0000 < X = 0.01$$

แสดงว่ามีนัยสำคัญจึงยอมรับ  $H_1$  ที่ว่าลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นเส้นตรง

B) ทดสอบว่าความสัมพันธ์ระหว่างสวัสดิการในการทำงานและความตั้งใจที่จะออกจากองค์การมีการเบี่ยงเบนออกจากเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสอง ไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเบี่ยงเบนออกจากความเป็นเส้นตรง

$$P = 0.4934 < X = 0.01$$

แสดงว่าไม่มีนัยสำคัญจึงยอมรับ  $H_0$  ที่ว่า ลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

#### ตารางที่ 26

แสดงผลของการวิเคราะห์ความแปรปรวนระหว่างความยึดมั่นผูกพันในองค์การกับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	3385.713	2	1692.856	67.748	0.0000
Linearity	3379.420	1	3379.420	135.245	0.0000
DEV. from linearity	6.293	1	6.293	0.252	0.6163

A) ทดสอบว่าความสัมพันธ์ระหว่างความยึดมั่นผูกพันในองค์การและความตั้งใจที่จะออกจากองค์การมีลักษณะเป็นเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสอง ไม่ได้เป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสอง เป็นเส้นตรง



$$P = 0.0000 < X = 0.01$$

แสดงว่ามีนัยสำคัญจึงยอมรับ  $H_1$  ที่ว่าลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นเส้นตรง

B) ทดสอบว่าความสัมพันธ์ระหว่างความยึดมั่นผูกพันในองค์การ และความตั้งใจที่จะออกจากองค์การมีการเบี่ยงเบนออกจากเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเบี่ยงเบนออกจากความเป็นเส้นตรง

$$P = 0.6163 < X = 0.01$$

แสดงว่าไม่มีนัยสำคัญจึงยอมรับ  $H_0$  ที่ว่า ลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

#### ตารางที่ 27

แสดงผลของการวิเคราะห์ความแปรปรวนระหว่างความพึงพอใจในการทำงาน  
กับความตั้งใจที่จะออกจากองค์การ

Source	S.S.	df.	M.S.	F.	Sig
Between groups	2312.900	2	1156.450	38.947	0.0000
Linearity	2308.827	1	2308.827	77.757	0.0000
DEV. from linearity	4.073	1	4.073	0.137	0.7115

A) ทดสอบว่าความสัมพันธ์ระหว่างความพึงพอใจในการทำงานและความตั้งใจที่จะออกจากองค์การมีลักษณะเป็นเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเป็นเส้นตรง

$$P = 0.0000 < X = 0.01$$

แสดงว่ามีนัยสำคัญจึงยอมรับ  $H_1$  ที่ว่าลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองเป็นเส้นตรง

B) ทดสอบว่าความสัมพันธ์ระหว่างความพึงพอใจในการทำงานและความตั้งใจที่จะออกจากองค์กรมีการเบี่ยงเบนออกจากเส้นตรงหรือไม่?

$H_0$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

$H_1$  : ลักษณะความสัมพันธ์ของตัวแปรทั้งสองเบี่ยงเบนออกจากความเป็นเส้นตรง

$$P = 0.7115 < X = 0.01$$

แสดงว่าไม่มีนัยสำคัญจึงยอมรับ  $H_0$  ที่ว่า ลักษณะความสัมพันธ์ระหว่างตัวแปรทั้งสองไม่ได้เบี่ยงเบนออกจากความเป็นเส้นตรง

ดังนั้น สำหรับฐานคติเบื้องต้นที่ว่า ความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามต้องมีลักษณะเป็นเส้นตรงถือว่าผ่านไป

2.2 จะต้องไม่มีตัวแปรอิสระใด ๆ ที่มีความสำคัญต่อการอธิบายตัวแปรตามอยู่นอกตัวแบบที่ใช้ในการศึกษา

2.3 จะต้องไม่มีตัวแปรอิสระใด ๆ ที่ไม่มีความสำคัญต่อการอธิบายตัวแปรตามอยู่ในตัวแบบที่ใช้ในการศึกษา

สำหรับฐานคติทั้ง 2 ข้อนี้ จะกล่าวรวมกันไปเพราะปัญหาทั้งสองอย่างนี้มีความใกล้เคียงกันมาก อีกทั้งถือว่ามีความสำคัญอย่างยิ่งในการสร้างตัวแบบเพื่อการศึกษาปรากฏการณ์ต่าง ๆ เพราะว่าการศึกษาจะปรากฏออกมาเช่นไรขึ้นอยู่กับ การคัดเลือกตัวแปร ปรากฏการณ์ด้านสังคมศาสตร์มีความซับซ้อนและมีเงื่อนไขมากมาย กล่าวคือ มีตัวแปรที่มีความสัมพันธ์กันหลากหลาย จนในบางครั้งก็ไม่ทราบเหมือนกันว่า ตัวแปรที่นำมาบรรจุในตัวแบบนั้นจะมีความสำคัญอย่างไรหรือไม่ต่อการอธิบายปรากฏการณ์ที่กำลังศึกษา และในกรณีเดียวกัน ตัวแปรที่ไม่ได้นำมาบรรจุในตัวแบบ เราก็ไม่ทราบเช่นเดียวกันว่า จะไม่มีผลต่อการอธิบายปรากฏการณ์อย่างจริง ๆ สำหรับปัญหาเกี่ยวกับฐานคติทั้งสองนี้ ไม่ว่าจะใช้สถิติประเภทใด ๆ ก็ไม่อาจจะ

ทดสอบได้ และเมื่อใดก็ตามที่ปัญหานี้บังเกิดขึ้น อำนาจการอธิบายของตัวแบบที่สร้างขึ้นจะน้อยลง และค่าความผิดพลาดจะสูงมากขึ้น (ประชัย เปี่ยมสมบูรณ์, 2527 : 74)

ในการตรวจสอบฐานคติเบื้องต้นลักษณะนี้ โดยทางปฏิบัติแล้วจะต้องอาศัยการตรวจสอบโดยแนวคิดเชิงทฤษฎีในแต่ละสาขาวิชาเป็นพื้นฐานสำคัญ (Lewis-beck, 1980 : 27) สำหรับในการศึกษาคั้งนี้ตัวแบบที่ได้สร้างขึ้นนั้น อาศัยพื้นฐานทางทฤษฎีและวรรณกรรมที่เกี่ยวข้องเป็นหลัก ซึ่งจะทำให้การละเมิดฐานคติทั้ง 2 ข้อนี้ ไม่รุนแรง และสามารถนำตัวแบบมาใช้ในการศึกษาได้

2.4 ค่าความแปรปรวนผิดพลาดจะต้องมีค่าคงที่ (homoscedasticity :

$$E(E^2) = b^2$$

ความคงที่ของค่าความแปรปรวนที่ไม่ถูกอธิบาย อาจเรียกอีกอย่างหนึ่งได้ว่า "ความกระจายคงที่" ซึ่งหมายความว่า ค่าความแปรปรวนที่เหลือจากการอธิบายโดยตัวแปรอิสระมีช่วงระยะของจุดข้อมูลคงที่สำหรับทุกค่าของตัวแปรอิสระ หากค่าความแปรปรวนดังกล่าวไม่คงที่จะเรียกว่า heteroscedasticity (ยุบลวรรณ ประมวลรัฐการ, 2532 : 266) การละเมิดฐานคตินี้ถือว่าเป็นปัญหาที่มีความรุนแรงมากที่สุด เพราะจะทำให้การทดสอบเกี่ยวกับสมมติฐานและช่วงของความเชื่อมั่นผิดพลาดได้ (Lewis-beck, 1980 : 28)

วิธีการสะดวกที่สุดในการพิจารณา คือ การเขียนกราฟ โดยการกำหนดค่าผิดพลาดกับค่าพยากรณ์ของตัวแปรตาม (residuals against the predicted values) ว่า ปรากฏออกมาในลักษณะใด หากว่ากราฟออกมาเป็นรูปแบบที่กระจายกัน (randomly distributed) คือ จุดต่าง ๆ ในรูปกราฟมีการกระจายเป็นรูปขนาน แสดงว่าค่าพยากรณ์จะไม่มีความสัมพันธ์กับค่าผิดพลาด คือ ไม่ว่าค่าพยากรณ์จะเปลี่ยนแปลงไปเท่าใดก็ตาม ค่าความผิดพลาดจะคงที่อยู่เสมอ (variance constant) หรืออาจจะรู้จักชื่อที่เรียกกันว่าลักษณะของความเท่าเทียมกันของความแปรปรวน (equality of variance) หรือค่าความแปรปรวนมีลักษณะเป็นอันหนึ่งอันเดียวกัน (homogeneity of variance) (Norusis, 1985: 24- 26 ; Seber, 1977 : 165) ในการตรวจสอบการกระจายตัวแปรตามและผลปรากฏออกมาดังนี้



1. ความผิดพลาดจากมโนทัศน์ เกี่ยวข้องกับข้อความที่ใช้วัดมโนทัศน์นั้นขาดความสอดคล้องกับมโนทัศน์ (ตัวแปร) ทำให้การวัดผิดพลาด ส่งผลถึงความลำเอียงในการประมาณค่าสัมประสิทธิ์ถดถอย

2. ความผิดพลาดด้านความสอดคล้องหรือความผิดพลาดเชิงระบบ ซึ่งเกิดจากเหตุผลหลายประการณ ดังเช่น ผู้ตอบแบบสอบถามให้ข้อมูลผิดพลาด (เช่น รายได้ อายุ) ความผิดพลาดจากเครื่องมือวัด ผู้สัมภาษณ์ เป็นต้น ความผิดพลาดประเภทนี้จะทำให้เกิดความลำเอียงในการประมาณค่าสัมประสิทธิ์ถดถอย ทิศทาง และขนาดของความลำเอียงไม่สามารถกำหนดได้ถ้าปราศจากการรับรู้ทิศทางและขนาดของความผิดพลาด

3. ความผิดพลาดเชิงสุ่มหรือไม่เป็นระบบ ซึ่งเป็นผลมาจากความแปรเปลี่ยนเกี่ยวกับเวลาในด้านของผู้ตอบแบบสอบถาม ผู้สัมภาษณ์ เป็นต้น

สำหรับการสร้างข้อความในการวัดครั้งนี้ ได้พยายามยึดถือผลงานวิจัยที่เกี่ยวข้องในอดีตและยังพิจารณาถึงความเหมาะสมขององค์การที่เข้าไปศึกษาด้วย ในการวัดตัวแปรอิสระคือ โอกาสความก้าวหน้าในการทำงาน ความเป็นอิสระในการทำงาน การสื่อสารในองค์การ การรวมอำนาจ รายได้ สวัสดิการในการทำงาน ความยึดมั่นผูกพันในองค์การ ความพึงพอใจในการทำงาน และตัวแปรตามก็คือ ความตั้งใจที่จะออกจากองค์การ แม้ว่าตัวแปรบางตัวจะกำหนดมาใหม่ เช่น สวัสดิการในการทำงาน แต่ผู้วิจัยก็ศึกษาจากระบบขององค์การที่เข้าไปศึกษา แล้วนำตัวแปรต่าง ๆ มาทดสอบความน่าเชื่อถือ ซึ่งผลก็อยู่ในเกณฑ์ดี ดังนั้น ในประเด็นเกี่ยวกับการวัดตัวแปรจะต้องไม่มีการผิดพลาดในด้านมโนทัศน์สำหรับการศึกษาคั้งนี้ ก็น่าจะอนุโลมให้ผ่านไปได้

ส่วนความผิดพลาดเชิงระบบและความผิดพลาดเชิงสุ่มนั้น หลีกเลี้ยงได้ยาก เพราะปัญหาเหล่านี้อาจเกิดขึ้นโดยไมทันได้ระวัง แต่อย่างไรก็ตาม ในการวิจัยครั้งนี้ผู้วิจัยก็ได้กระทำการอย่างรอบคอบและประเมินว่า ความผิดพลาดด้านนี้คงจะมีไม่มากนัก ซึ่งก็น่าจะส่งผลให้ได้ข้อมูลตามความเป็นจริงมากที่สุด

### 3.1 ปัญหาการมีสหสัมพันธ์ระหว่างตัวแปรอิสระ (multicollinearity)

การทดสอบฐานคติเกี่ยวกับสหสัมพันธ์ระหว่างตัวแปรอิสระก็เป็นประเด็นสำคัญอันหนึ่งสำหรับการวิเคราะห์ถดถอย การมีสหสัมพันธ์ระหว่างตัวแปรอิสระจะไม่พบปัญหาสำหรับการวิเคราะห์ถดถอยแบบง่าย แต่ปัญหานี้มักจะเกิดขึ้นกับการวิเคราะห์ถดถอยแบบพหุ ถ้า

ตัวแปรอิสระมีความสัมพันธ์กันอย่างสูง จะทำให้เพิ่มอำนาจในการอธิบายได้ต่ำ เพราะว่าผลกระทบของตัวแปรอิสระตัวหนึ่งขึ้นอยู่กับค่าตัวแปรอิสระอื่น ๆ ในการที่จะอธิบายตัวแปรตาม รวมทั้งยังส่งผลให้เกิดความยากลำบากในการเปรียบเทียบ และแยกแยะความสำคัญของตัวแปรอิสระ หรือผลกระทบของตัวแปรอิสระทั้งหลายที่มีต่อตัวแปรตาม และทำให้การตีความค่าสหสัมพันธ์ในเชิงสาเหตุและผลเป็นไปอย่างซับซ้อน (ปุระชัย เปี่ยมสมบูรณ์, 2524 : 41-45)

ดังนั้น ก่อนนำตัวแปรอิสระต่าง ๆ เข้าไปวิเคราะห์ในสมการถดถอยพหุ จะต้องนำค่าสหสัมพันธ์ของตัวแปรอิสระมาพิจารณาเสียก่อน ในที่นี้จะแสดงตารางที่ 28

ตารางที่ 28

แสดงค่าสหสัมพันธ์ของตัวแปรอิสระ

ตัวแปร	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$
$X_1$	1.000							
$X_2$	.276	1.000						
$X_3$	.210	.432	1.000					
$X_4$	-.176	-.432	-.384	1.000				
$X_5$	.035	.091	.094	-.060	1.000			
$X_6$	.265	.258	.276	-.264	.088	1.000		
$X_7$	.242	.321	.351	-.271	.204	.492	1.000	
$X_8$	.295	.351	-.310	-.310	.170	.532	.644	1.000

จากค่าดังกล่าวก็คงสรุปได้ว่า ปัญหาสหสัมพันธ์ระหว่างตัวแปรอิสระ จะไม่มีผลกระทบในการศึกษาครั้งนี้ เพราะตัวแปรส่วนใหญ่มีค่าสหสัมพันธ์ค่อนข้างต่ำ สำหรับฐานคติข้อนี้ก็น่าจะผ่านไปได้

#### 4. ฐานคติเบื้องต้นในประเด็นเกี่ยวกับค่าผิดพลาด (error term, $E_1$ )

ประเด็นที่จะนำมาพิจารณาต่อไปนี้จะเป็นเรื่องของการวิเคราะห์ผิดพลาด

##### 4.1 ค่าผิดพลาด (residuals) จะต้องมีค่าเฉลี่ยเท่ากับ 0 ( $E(E_1) =$

0) ประเด็นดังกล่าวก็คือ ค่าความคาดหวัง (expected value) ของค่าผิดพลาดจะมีค่าเป็น 0 นั่นเอง เพราะว่าสำหรับตัวแปรที่มีการกระจายอย่างปกติแล้ว ค่าความคาดหวังจะมีค่าเท่ากับค่าเฉลี่ยของตัวมันเอง ซึ่งในที่นี้ก็คือค่า 0 ดังกล่าว (Lewis-beck, 1980 : 27) ถ้าหากพบว่าค่าไม่เท่ากับ 0 แล้วจะทำให้ค่าการประเมินลักษณะของประชากรเกิดความไม่เที่ยงตรงสำหรับการวิเคราะห์ถดถอย แล้วจำเป็นที่จะต้องใช้หลักกำลังสองต่ำสุด ดังนั้น ค่าความผิดพลาดก็จะมีค่าเท่ากับ 0 อยู่แล้ว

##### 4.2 ตัวแปรอิสระจะต้องไม่มีความสัมพันธ์กับค่าผิดพลาด : $E(E_1 X_1) = 0$

ฐานคติเบื้องต้นเกี่ยวกับตัวแปรอิสระต้องไม่มีความสัมพันธ์กับค่าผิดพลาด นั้น จะพบได้ยากมากที่เดียวสำหรับการวิจัยที่มีการทดลอง และเมื่อใดก็ตามที่ตัวแปรอิสระได้ไม่มีความสัมพันธ์กับค่าผิดพลาดแล้ว จะส่งผลให้การประมาณค่าของหลักกำลังสองต่ำสุดเกิดความคลาดเคลื่อนได้ (Lewis-beck, 1980 : 29) สำหรับการตรวจสอบจะใช้หลักเกณฑ์พิจารณาอันเดียวกันกับการทดสอบรูปแบบของค่าความแปรปรวนผิดพลาดต้องมีค่าคงที่ (homoskedasticity) เพียงแต่ค่าของตัวแปรอิสระทุกตัวซึ่งได้รับการปรับมาแล้วจะเข้าไปแทนที่ค่าพยากรณ์เท่านั้นเอง และค่าผิดพลาดมาตรฐานยังคงเดิมอยู่ในสมการ หากว่าจุดต่าง ๆ ในกราฟที่ปรากฏออกมา มีการกระจายกันอย่างปกติ แสดงว่าตัวแปรอิสระแต่ละตัวกับค่าผิดพลาดไม่มีความสัมพันธ์กัน ซึ่งฐานคติเบื้องต้นข้อนี้จะเป็นข้อหนึ่งสำหรับการวัดความสมบูรณ์ของตัวแปรอิสระที่จะนำมาใช้วิเคราะห์ในสมการถดถอย หากผ่านการทดสอบข้อนี้ไปก็เท่ากับว่าผ่านไปอีกหนึ่งข้อ ซึ่งจะมีผลดีต่อการวิเคราะห์มากยิ่งขึ้น

การพิจารณาจากรูปกราฟสามารถดูได้จากภาพที่ 15-22 ซึ่งจากการพิจารณารูปกราฟแล้วพบว่า ค่าความแปรปรวนผิดพลาดมีค่าคงที่ ดังนั้น ฐานคติข้อนี้ก็ผ่านการตรวจสอบ











4.3 ค่าความผิดพลาดต้องกระจายอย่างปกติ (error term is normally distributed)

ลักษณะการกระจายของค่าผิดพลาด จะต้องมีการกระจายกันอย่างปกติ การตรวจสอบฐานคติข้อนี้ จะกระทำกันโดย เขียนกราฟของค่าความผิดพลาดมาตรฐานแจกแจงออกมาในรูปของกราฟแท่งและแบบโค้งปกติควบคู่กัน ไปดังภาพที่

ปกติแล้วการเขียนกราฟของค่าผิดพลาดมาตรฐาน มีประโยชน์ในการค้นคว้าหาว่ามีค่าผิดพลาดมาตรฐานอื่น ๆ อีกหรือไม่ที่กระจายออกจากโค้งปกติมากกว่าค่ากระจายอย่างปกติแล้ว จะพบว่าประมาณ 68% ของค่าผิดพลาดมาตรฐานจะตกอยู่ในช่วง  $-1$  ถึง  $1$  และประมาณ 95% ของค่าดังกล่าวจะตกอยู่ในช่วง  $-2$  ถึง  $+2$  จากภาพที่ ปรากฏว่าการกระจายของค่าผิดพลาดส่วนใหญ่จะอยู่ใน โค้งปกติและค่าผิดพลาดมาตรฐานทั้งหมดจะอยู่ระหว่าง  $-2.41644$  (minimum standardized residuals) และ  $+1.90113$  (maximum standardized residuals) แต่ประเด็นที่น่าสนใจก็คือ ค่าผิดพลาดที่เรียกว่า large residuals หรือ (outliers) ซึ่งเป็นค่าผิดพลาดที่กระเด็นห่างออกไปจากค่าเฉลี่ยของความผิดพลาดมาตรฐานจนเกินไป หรือเป็นค่าที่ตกอยู่นอกขอบเขต  $\pm 2$  หรืออาจจะเป็นค่าผิดพลาดซึ่งมีค่าเบี่ยงเบน มาตรฐานมากกว่า 1 (ประสิทธิ์ สันติกาญจน์, 2530 : 148)

จากการวิเคราะห์พบว่าไม่มี outliers จึงคงข้อมูลไว้โดยไม่ต้องตัดชุดใดชุดหนึ่งทิ้ง จึงถือว่าการวิเคราะห์ประเด็น outliers ที่จะมีผลกระทบต่อ การวิเคราะห์ถดถอยหุ้้น เป็นอันว่าผ่านขั้นตอนนี้ไป

แต่ปัญหาที่จะตามมาอีกประการหนึ่ง ก็คือ ส่วนที่เป็นค่าผิดพลาดซึ่งเท่าที่ปรากฏอยู่หุ้้น ถ้าพิจารณาในระดับรวม ๆ กันแล้ว (gross normality) อาจมีผลกระทบต่อค่า  $t$  หรือค่า  $F$  หรืออาจมีผลต่อช่วงของความเชื่อมั่นและช่วงของการทำนาย (prediction intervals) ก็เป็นได้ จึงต้องเขียนกราฟในรูปแบบที่เรียกว่า "normal probability plots" เป็นการพิจารณาค่าผิดพลาดจากข้อมูลดิบที่สำรวจมาได้ว่ามีโอกาสจะเป็นเส้นตรงหรือไม่ โดยที่จะกำหนดค่าผิดพลาดทั้งหมดลงไปในรูปแบบของ  $e_1 < e_2 < e_3 \dots < e_n$  เรียงลำดับสูงขึ้นไปเรื่อย ๆ ควบคู่กันกับการกระจายอย่าง เป็นปกติที่เป็นไปได้ในลักษณะสะสม (cumulative probability normal-distribution) ซึ่งจะมีลักษณะของจุดต่าง ๆ ที่เรียงกันเป็นเส้นตรง จากนั้นจะพิจารณาค่าความผิดพลาดมีรูปแบบใกล้เคียงกันกับเส้นตรง

ดังกล่าวหรือไม่ (Montgomery and Pick, 1982 : 59) และพิจารณาจากกราฟแบบ normal probability plots ดังภาพที่ 24

จากการพิจารณารูปกราฟพบว่า มีความใกล้เคียงกับความเป็นเส้นตรง จึงสรุปได้ว่า การกระจายของค่าผิดพลาดมาตรฐานนี้จะมีลักษณะเป็นปกติ เพราะไม่ว่าส่วนหัวหรือส่วนท้ายของเส้นกราฟไม่มีความเยื้องหรือความโค้งอย่างเด่นชัด สำหรับฐานคติข้อนี้ก็ถือว่าผ่านได้

ภาพที่ 23

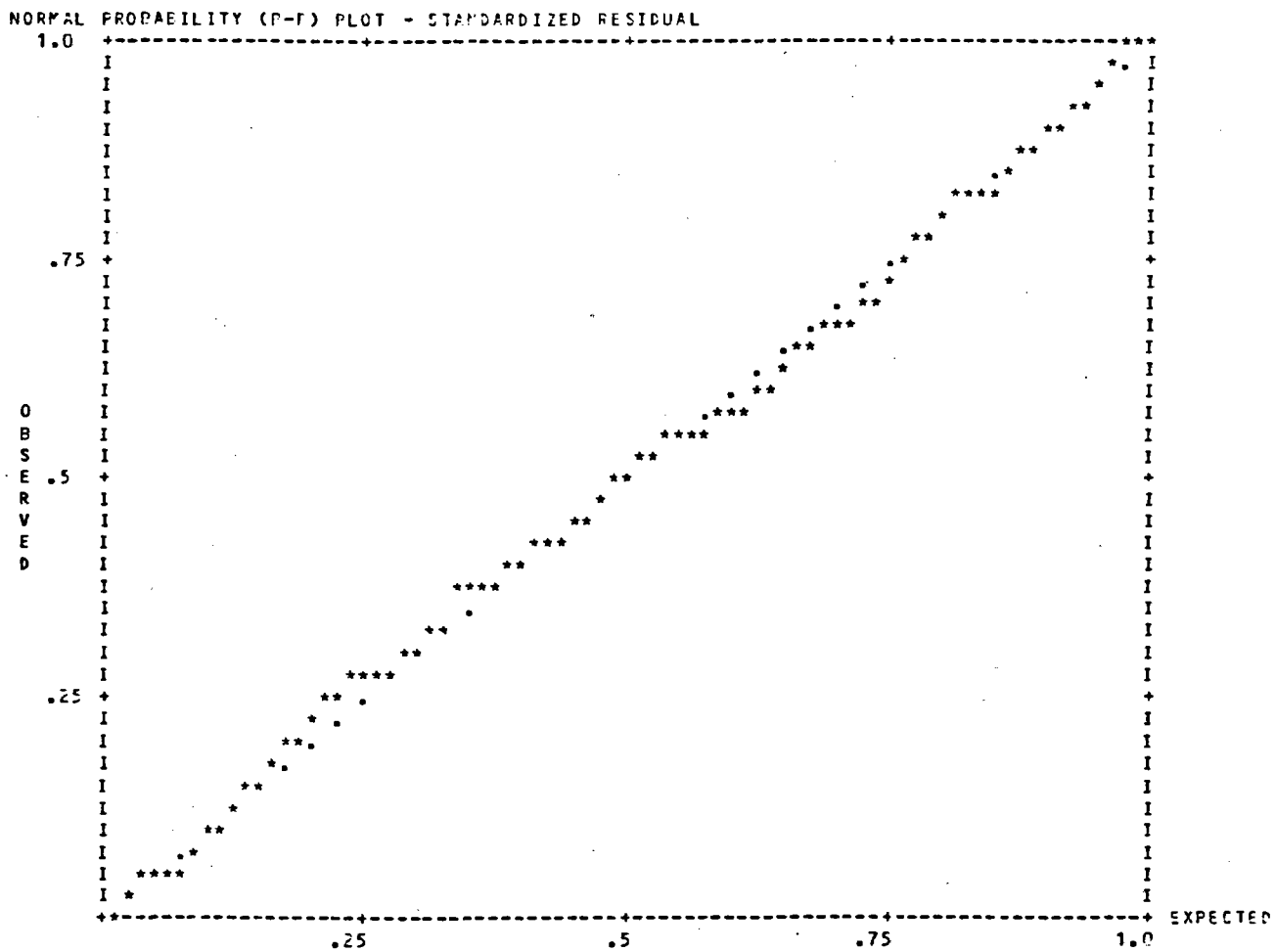
แสดงการกระจายของค่าผิดปกติมาตรฐาน

HISTOGRAM - STANDARDIZED RESIDUAL  
 (\* = 1 CASES, . = NORMAL CURVE)

N	EXP	N	RESIDUAL	MARKERS
0	.25	OUT		
1	.13	3.00	*	
0	.19	2.88		
0	.26	2.75		
0	.37	2.63		
0	.51	2.50	.	
0	.69	2.38	.	
0	.92	2.25	.	
0	1.21	2.13	.	
1	1.56	2.00	*	
1	1.99	1.88	*	
6	2.49	1.75	*:****	
5	3.08	1.63	**:*	
5	3.74	1.50	***:*	
3	4.48	1.38	***.	
6	5.28	1.25	****:*	
8	6.12	1.13	*****:*	
4	6.99	1.00	****.	
10	7.85	.88	*****:*	
9	8.69	.75	*****:	
15	9.47	.63	*****:*****	
7	10.16	.50	*****.	
14	10.73	.38	*****:***	
8	11.16	.25	*****.	
7	11.42	.13	*****.	
10	11.51	0.0	*****.	
11	11.42	-.13	*****:	
12	11.16	-.25	*****:*	
7	10.73	-.38	*****.	
12	10.16	-.50	*****:*	
7	9.47	-.63	*****.	
8	8.69	-.75	*****.	
11	7.85	-.88	*****:*	
9	6.99	-1.00	*****:*	
8	6.12	-1.13	*****:*	
8	5.28	-1.25	*****:*	
3	4.48	-1.38	***.	
2	3.74	-1.50	**.	
3	3.08	-1.63	**:	
1	2.49	-1.75	*.	
3	1.99	-1.88	*:*	
1	1.56	-2.00	*.	
2	1.21	-2.13	:*	
2	.92	-2.25	:*	
1	.69	-2.38	:	
0	.51	-2.50	.	
0	.37	-2.63		
0	.26	-2.75		
0	.19	-2.88		
0	.13	-3.00		
0	.25	OUT		

ภาพที่ 24

แสดงการกระจายของค่าผิดพลาดมาตรฐานในลักษณะของ Normal Probability Plot



### วิธีการแก้ไขการละเมิดฐานคติเบื้องต้น

1) การละเมิดฐานคติ Homoscedasticity ต้องใช้วิธีวิเคราะห์ที่เรียกว่า วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก (weighted least squares) กล่าวคือ หลังจากการตรวจสอบ residuals แล้ว ถ้าปรากฏว่า

$$\text{var}(E_1) = \sigma^2_1 = \sigma^2 \text{ สำหรับทุกค่าของ } X_1$$

เรียกว่าเกิดปัญหา Heteroscedasticity ซึ่งขัดกับฐานคติของการวิเคราะห์ถดถอย ทำให้ไม่สามารถใช้ OLS ในการประมาณค่าของตัวแบบได้ เนื่องจากค่าประมาณที่ได้จาก OLS กรณีนี้จะมีคุณสมบัติที่จะไม่ให้ค่าความแปรปรวนต่ำสุด แม้จะยังมีคุณสมบัติของความไม่เอนเอียงอยู่ก็ตาม ซึ่งจะมีผลต่อการทดสอบและการพยากรณ์

วิธีการของ WLS นี้คือ เราต้องการจะทำให้

$$s_1 = \frac{1}{\sigma_1^2} (Y_1 - b_0 - b_1 X_1)^2 \text{ มีค่าต่ำที่สุด} \quad (1)$$

$$\text{เมื่อตัวแบบที่เหมาะสมเป็น } Y_1 = \beta_0 + \beta_1 X_1 + E_1 \quad (2)$$

ในที่นี้สมมติว่า  $\sigma^2_1$  จะมีค่ามากขึ้นตามค่าของ  $X_1$  ที่มากขึ้น ทำให้  $\frac{1}{\sigma_1^2}$  มีค่าน้อยลง

ถ้าสมมติให้  $\sigma_1$  เพิ่มขึ้นเป็นสัดส่วนกับ  $X_1$  กล่าวคือ

$$\sigma_1 = kX_1 ; k > 0$$

$$\text{หรือ } \frac{\sigma_1}{X_1} = k$$

ดังนั้น การทำให้สมการ (1) มีค่าต่ำสุด ก็คือ การทำให้

$$\sum_{i=1}^n \frac{1}{k^2 X_1^2} (Y_1 - b_0 - b_1 X_1)^2 \text{ มีค่าต่ำสุดนั่นเอง}$$

ถ้าเปลี่ยนรูปตัวแบบใหม่ (transform equation โดยการเอา  $X_1$  หารสมการ

(2) ตลอด จะได้สมการใหม่เป็น

$$\frac{Y_1}{X_1} = \beta_0 \cdot \frac{1}{X_1} + \beta_1 + \frac{E_1}{X_1} \quad (3)$$



พิจารณา  $\text{Var} \left( \frac{E_{1-}}{x_{1-}} \right) = \frac{1}{x_{1-}^2} \text{Var}(E_1) = \frac{\sigma^2}{x_{1-}^2} = \frac{k^2 X^2}{x_{1-}^2} = k^2$  ซึ่งมีค่าคงที่ นั่นคือ เราสามารถใช้ OLS ในการประมาณค่าพารามิเตอร์จากตัวแบบที่เปลี่ยนรูปแล้ว ในสมการ (3)

2) การละเมิดฐานคติ Multicollinearity มีวิธีลดความรุนแรงของปัญหาโดยวิธีการดังต่อไปนี้

2.1 การตัดตัวแปร (dropping variables) ในทางปฏิบัตินิยมตัดตัวแปรที่มีความสำคัญน้อยกว่า หรือตัวแปรที่เราไม่สนใจทิ้งไป

2.2 การเพิ่มขนาดของตัวอย่าง (increase of the sample size) เป็นวิธีหนึ่งนี้อาจหลีกเลี่ยง หรือลดความรุนแรงของการเกิด Multicollinearity โดยการเพิ่มค่าสังเกตมากขึ้น

2.3 การใช้ตัวแปรล่าหลัง (lagged variables) เป็นตัวแปรอิสระอื่น ๆ ใน Distributed-Lag models

2.4 การใช้ ratios หรือ first differences ส่วนมากจะใช้ใน time-series analysis

3) วิธีการแปลงข้อมูล (transformations) จะใช้กับการละเมิดฐานคติต่อไปนี้

3.1 ฟังก์ชันการถดถอยไม่เป็นเส้นตรง (linearity)

การตรวจสอบจากความแปรปรวนระหว่างตัวแปรอิสระและตัวแปรตาม ทำให้ทราบว่า ตัวแปรมีความสัมพันธ์กันเป็นเส้นตรงหรือไม่ ถ้าพบว่าเป็น เราสามารถแปลงข้อมูลให้เป็นเส้นตรงได้ เรียกว่า intrinsically Linear การแปลงข้อมูลให้เป็นเส้นตรง จะใช้วิธีการ take log เข้าไปในสมการ เช่น

$$Y = \beta_0 + \beta_1 \log x \quad X' = \text{Log } x \quad Y' = \beta_0 + \beta_1 X'$$

3.2 การแปลงข้อมูลเพื่อให้เกิดค่าความแปรปรวนคงที่

ในกรณีนี้พบว่า ความแปรปรวนของ E มีค่าไม่คงที่ และมีความสัมพันธ์กับค่า E(Y) แล้ว สามารถใช้การแปลงข้อมูลเพื่อให้ความแปรปรวนมีค่าคงที่ได้ ดังกรณีเช่น

ถ้า Y มีการแจกแจงแบบพิวของแล้ว ค่าความแปรปรวนของ Y จะมีค่าเท่ากับค่าเฉลี่ย Y เนื่องจาก Y มีความสัมพันธ์กับตัวแปรอิสระ X ดังนั้น ความแปรปรวนของ

$Y$  จึงเป็นส่วนเดียวกับ  $X$  ในกรณีนี้ควรใช้วิธีการ squar root กล่าวคือ  $Y' = \sqrt{Y}$  จะทำให้ความแปรปรวนของ  $\sqrt{Y}$  เป็นอิสระกับค่าเฉลี่ย

ถ้า  $Y$  มีค่าเป็นส่วน (  $0 < Y < 1$  ) และ residual plots ของ  $e_1$  กับ  $Y_1$  แสดงให้เห็น double bow แล้ว ควรใช้การแปลงให้อยู่ในรูปของ arcsin คือ  $Y' = 2 \arcsin \sqrt{Y}$

จากที่กล่าวมาทั้งหมดนั้น คือ วิธีการที่จะแก้ไขการละเมิดฐานคติของการวิเคราะห์ถดถอยแบบพหุ ซึ่งเป็นการกล่าวถึงแนวทางโดยสังเขป สำหรับรายละเอียดต้องค้นคว้าเพิ่มเติมจากหนังสือสถิติทั่วไป