บทที่ 2

วิธีการแก้ปัญหา

วิธีการที่ใช้ในการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่า และจำนวนครั้งที่เรียกโทรศัพท์ในเขตนครหลวงมีหลายวิธี แต่ในที่นี้จะกล่าวถึงเพียง 2 วิธี คือ

1. <u>Trend Method</u> วิธีนี้ให้ตัวแปร (Variable) เป็น function ของเวลา คือ

$$Y = f(X) + e$$

Y = ตัวแปร (Variable)

X = เวลา (Time)

โมเดลที่ใช้ในการคำนวณหาเส้นแนวโน้ม (Trend Line) มี

ก. Linear Trend เส้นแนวโน้มที่เป็นเส้นตรงมีรูปสมการเป็น

$$Y = a + bX + e \qquad (1)$$

Y = ตัวแปร (Variable)

X = เวลา (Time)

a = intercept ของ Y

b = trend increment หรือ slope ของ Y

วิธีหาค่า a,b ใช้วิธีการของ least squares คือให้ผลรวมของความแตกต่างของจุด Y ยกกำลังสองมีค่าน้อยที่สุด จะให้ normal equation ดังนี้

$$\sum Y = na + b\sum X$$
$$\sum XY = a\sum X + b\sum X^2 \qquad (2)$$

จากสมการ (2) จะ solve หาค่า a,b ได้

ข. Non-Linear Trend เส้นแนวโน้มที่ไม่เป็นเส้นตรง มีรูปสมการหลายรูปดังนี้

1) Second-degree polynomial curve เส้นแนวโน้มที่มีลักษณะเป็น parabola มีรูปสมการเป็น

$$Y = a + bX + cX^2 + e \qquad (3)$$

Y = ตัวแปร (Variable)

X = เวลา (Time)

a = intercept ของ Y

b = slope ของ Y

c = rate of change of slope

วิธีหาค่า a,b,c ใช้วิธีการของ least squares เช่นกัน ซึ่งให้ normal equation ดังนี้

$$\begin{aligned} \Sigma Y &= na + b\Sigma X + c\Sigma X^2 \\ \Sigma XY &= a\Sigma X + b\Sigma X^2 + c\Sigma X^3 \\ \Sigma X^2Y &= a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4 \end{aligned} \qquad (4)$$

จากสมการ (4) นี้ solve หาค่า a,b,c ได้

2) Simple Exponential Curve มีรูปสมการเป็น

$$Y = ab^X + e \qquad (5)$$

a = intercept

b = rate of increase

ถ้า b มีค่าระหว่าง 0 กับ 1 แสดงว่าค่า Y จะลดลง เมื่อ X มีค่าเพิ่มขึ้น

ถ้า b มีค่ามากกว่า 1 แสดงว่าค่า Y จะเพิ่มขึ้นเมื่อ X มีค่าเพิ่มขึ้น

จากรูปสมการ (5) เมื่อ take logarithm ทั้งสองข้างแล้วจะ transform อยู่ในรูปเส้นตรงได้คือ

$$\log Y = \log (ab^X) = \log a + X \log b$$

ดังนั้นการหาค่า a,b จึงใช้วิธี least squares ได้ ซึ่งจะให้ normal equation ดังนี้

$$\Sigma \log Y = n \log a + (\log b)\Sigma X$$
$$\Sigma X \log Y = (\log a)\Sigma X + (\log b)\Sigma X^2 \quad (6)$$

จากสมการ (6) solve หาค่า log a, log b ได้

3) Modified Exponential Curve มีรูปสมการเป็น

$$Y = L + ab^X \quad (7)$$

ลักษณะของเส้นแนวโน้มคล้ายๆกับ Simple Exponential Curve เพียงแต่บวกค่าคงที่ L เข้าไปอีก 1 ค่า ซึ่งทำให้ curve shift ขึ้นหรือลงด้วยค่าคงที่นั้น การที่ curve shift นี้เป็นการบอกค่าของ curve ว่ามีขีดจำกัดสูงหรือต่ำ

วิธีแก้สมการ ใช้วิธีการของ least squares ไม่ได้ เพราะเราไม่สามารถจะแปลงรูปสมการให้อยู่ในรูปของเส้นตรงอย่างเช่นใน Simple Exponential Curve ได้ ดังนั้นวิธีหาค่า L,a,b จะหาโดย Method of selected points[1]

4) Logistic Curve มีรูปสมการเป็น

$$Y = \frac{1}{L+ab^X} \quad (8)$$

Curve นี้อัตราการเพิ่มขึ้นหรือลดลงจะเป็นไปอย่างรวดเร็วในตอนแรก แต่เมื่อผ่านจุดๆหนึ่งแล้วจะลดลงอย่างช้าๆ จะเห็นว่า Logistic curve เป็นส่วนกลับของ Modified Exponential Curve ดังนั้นการหาค่า L,a,b จึงทำเช่นเดียวกับใน Modified Exponential Curve

5) Gompertz Curve มีรูปสมการเป็น

$$Y = La^{b^X} \quad (9)$$

ถ้าค่าของ b อยู่ระหว่าง 0 กับ 1 b^X จะเข้าใกล้ 0 เมื่อ X มีค่าเพิ่มขึ้น ดังนั้นค่าของ Y จะเข้าใกล้ L

ถ้าค่าของ b มากกว่า 1 และค่าของ a มากกว่า 1 Curve จะเพิ่มขึ้นโดยไม่มีขอบเขต แต่ถ้าค่าของ a อยู่ระหว่าง 0 กับ 1 Curve จะเข้าใกล้ 0 ในขณะที่ X มีค่าเพิ่มขึ้น

---

[1] ดูรายละเอียดในหนังสือ Statistics for Economics โดย William I. Greenwald หน้า 222

สำหรับ Curve นี้สามารถแปลง (Transform) ให้อยู่ในรูป Modified Exponential Curve ได้โดยการ take logarithm ทั้ง 2 ข้างดังนี้

$$\log Y = \log L + b^{x}\log a$$

ให้ $\log Y = Y'$, $\log L = L'$, $\log a = a'$ สมการจะอยู่ในรูป

$$Y' = L' + a'b^{x}$$

ดังนั้นวิธีหาค่า unknown parameters จึงใช้วิธีการเดียวกันใน Modified Exponential Curve

ปัญหาในการตัดสินใจว่าจะใช้สมการใดในการประมาณหาเส้นแนวโน้มเป็นเรื่องที่ยุ่งยากลำบาก เพราะเมื่อเราใช้ค่า standard error of estimate ซึ่งหาได้จากสูตร

$$S = \sqrt{\frac{\sum(y-\hat{y})^2}{n-k}} \qquad (10)$$

Y = ค่าของตัวแปรไม่อิสระ (dependent variable)

$\hat{Y}$ = ค่าประมาณของตัวแปรไม่อิสระจากสมการ

n = จำนวน observation

k = จำนวน parameter

และเปอร์เซ็นต์ของการกระจายของตัวแปรที่ไม่อิสระ (dependent variable) ซึ่งคำนวนจาก

$$\frac{\sum(\hat{y}-\bar{\hat{y}})^2}{\sum(y-\bar{y})^2} \times 100 \qquad (11)$$

$\bar{Y}$ = ค่ามัชฌิมเลขคณิตของตัวแปรไม่อิสระ

$\bar{\hat{Y}}$ = ค่ามัชฌิมเลขคณิตของค่าประมาณตัวแปรไม่อิสระ

เป็นเครื่องมือในการตัดสินใจ โดยดูว่า ค่า standard error of estimate และเปอร์เซ็นต์ของการกระจายของตัวแปรไม่อิสระของแต่ละสมการนั้น สมการใดให้ค่า standard error of estimate น้อยที่สุด และมีเปอร์เซ็นต์ของการกระจาย

ของตัวแปรไม่อิสระมากที่สุด ก็จะใช้สมการนั้นในการประมาณ[2] แต่บางทีมีปัญหาเกิดขึ้น คือ เมื่อเราตัดสินใจตามที่กล่าวมาแล้วนั้น สมการที่นำมาพยากรณ์ค่าในอนาคตจะมีแนวโน้มที่ diverge ค่อนข้างเร็ว ทำให้ค่าในอนาคตเปลี่ยนแปลงอย่างรวดเร็วเกินความเป็นจริง ไม่เป็นไปตามสภาพของข้อมูล ดังนั้นบางครั้งเราจึงต้องเลือกสมการอื่นที่เหมาะสม และสามารถให้เหตุผลพอในการประมาณ และการพยากรณ์ด้วย[3]

2. <u>Multiple Linear Regression Method</u> วิธีนี้เป็นการประมาณ โดยพิจารณาปัจจัยต่างๆที่เกี่ยวข้องด้วย กล่าวคือ ศึกษาความสัมพันธ์ในรูปสมการกำลังหนึ่ง

$$Y = a + b_1X_1 + b_2X_2 + \ldots\ldots + b_kX_k$$

Y = ตัวแปรไม่อิสระ (dependent variable)

X_i = ตัวแปรอิสระ (independent variable) $i = 1,2,\ldots, k$

b_i = partial regression coefficient $i = 1,2,\ldots,k$

a = intercept ของสมการ

ทั้งนี้ใช้ข้อมูลในอดีตคำนวนหา regression function ในรูปกำลังหนึ่ง ที่จะให้ best possible fit ระหว่าง Y กับ X_i

วิธีการสร้าง regression function ใช้หลักการของ Stepwise multiple regression โดยการเพิ่มค่า X ทีละตัวในสมการ ตามลำดับความสำคัญ หรือความสัมพันธ์มากน้อยระหว่าง X_i กับ Y ตัวแปรอิสระ (independent variable) ตัวแรกที่จะนำเข้ามาในสมการ regression จะต้องเป็นตัวที่ให้ค่ากำลังสองของสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) ระหว่าง X_i กับ Y สูงสุด สมมติให้เป็น X_1 นั่นคือ ให้ $V_1 = r^2_{yx_1}$ สูงสุด ดังนั้น สมการจะเป็น

$$Y = a + b_1X_1$$

---

[2] K.W. Smillie, <u>An Introduction to Regression and Correlation</u>, (Toronto: The Ryerson Press, 1966) p. 81

[3] Donald L, Harnett, <u>Introduction to Statistical Methods</u>, (Addison-Wesley Publishing Company, 1970) p.375

ซึ่งมี Residual Sum of Squares (SSR) = $SSR_1 = (V_1)$ SST โดยที่ SST คือ Sum of Squares Total ซึ่งเท่ากับ $\Sigma(y - \bar{y})^2$

ตัวแปรอิสระตัวต่อไปที่จะนำเข้ามาในสมการ regreesion จะต้องเป็นตัวที่ให้สัดส่วนของการกระจายที่เหลือสูงสุด สมมติให้เป็น X_2 นั่นคือ

$$V_2 = (1-r^2_{yx_1})\, r^2_{yx_2 \cdot x_1}$$

สูงสุดหมายความว่าสัดส่วนของการกระจายทั้งหมดของตัวแปรไม่อิสระที่เกิดจาก X_2 หลังจากการกระจายที่เกิดจาก X_1 แล้ว ดังนั้นสมการจะเป็น

$$Y = a' + b_1' X_1 + b_2 X_2$$

ซึ่งมี $SSR_2 = (V_1 + V_2)SST$

ตัวแปรอิสระตัวต่อไปที่จะนำเข้ามาในสมการ regression สมมติ X_3 จะต้องเป็นตัวที่ให้ค่า

$$V_3 = (1 - r^2 yx_1)(1 - r^2 yx_2 \cdot x_1) r^2 yx_3 \cdot x_1 x_2$$

สูงสุด นั่นคือ variable ที่ให้ค่าสัดส่วนของการกระจายทั้งหมดของตัวแปรไม่อิสระที่เกิดจาก X_3 สูงสุด หลังจากการกระจายที่เกิดจาก X_1, X_2 แล้ว ดังนั้นสมการจะเป็น

$$Y = a'' + b_1'' X_1 + b_2' X_2 + b_3 X_3$$

ซึ่งมี $SSR_3 = (V_1 + V_2 + V_3)SST$

ทำเช่นนี้ไปเรื่อยๆ สมมติมี X k ตัว จะได้สมการ regression เป็น

$$Y = a^{(k)} + b_1^{(k)} X_1 + b_2^{(k)} X_2 + \ldots\ldots + b_k X_k$$

ซึ่งมี $SSR_k = (V_1 + V_2 + \ldots\ldots + V_k)SST$

การพิจารณาเลือกว่าจะใช้สมการ regression ที่มี X กี่ตัว หมายความว่าค่า X ทั้ง k ตัวนั้น บางตัวไม่มีความสำคัญ (significant) พอที่ จะใส่ในสมการโดยควรจะรวมอยู่ในค่า Residual ดังนั้นจะตัดค่า X ที่ ไม่มีความสำคัญออกไม่ต้องใส่ในสมการ เพราะถ้าใส่เข้าไปก็มิได้ทำให้ค่าประมาณ $\hat{Y}$ ใกล้เคียงกับค่าจริง Y มากขึ้น

วิธีการพิจารณา คือ ดูว่า standard error of estimate ของสมการในขั้นการเติม X_i ขั้นใด ให้ค่า standard error of estimate น้อยที่สุดแสดงว่าสมการ regression ที่มี X p ตัวโดยที่ $p < k$ ให้ค่าประมาณ $\hat{Y}$ ใกล้เคียงกับค่าจริงที่สุด และค่า F ในการทำ analysis of variance มากที่สุด แสดงว่า X จำนวน p ตัวที่อยู่ในสมการ regression มีอิทธิพลต่อ Y มากพอ จึงควรที่จะนำเข้าเป็นตัวแปรอิสระในสมการ[4]

2.1 การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าในเขตนครหลวง

2.1.1 โดย Trend Method แบ่งการประมาณตามประเภทผู้เช่าดังนี้

1) การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่ารวมทุกประเภท หมายถึงเป็นการประมาณจำนวนรวม ไม่คำนึงถึงประเภทต่างๆอันได้แก่ ประเภทร้านค้าบ้านพัก พิเศษ ราชการ สาธารณะ ทศท. จากข้อมูลที่เก็บรวบรวมไว้เป็นรายปีตั้งแต่ปี 2507 จนถึงปี 2515 ดังตารางที่ 2.1.1

Y_1 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่ารวมทุกประเภท

---

[4] K.W. Smillie, op. cit., p. 64

ตารางที่ 2.1.1

แสดงข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าแยกประเภทต่าง ๆ ตั้งแต่ปี 2507 - 2515

(พันเลขหมาย)

ปี	Z_1	Z_2	Z_3	Z_4	Z_5	Y'_1	Z_6	Y_1
2507	15.077	14.727	.926	3.809	.245	34.784	.280	35.064
2508	16.844	16.398	1.017	4.316	.285	38.860	.295	39.155
2509	18.391	17.179	1.207	4.593	.293	41.663	.397	42.060
2510	19.572	17.671	1.178	4.548	.319	43.288	.573	43.861
2511	22.566	18.741	1.279	4.859	.382	47.827	.343	48.170
2512	26.859	21.309	1.453	5.990	.436	56.047	.348	56.395
2513	29.659	25.982	2.423	7.254	.504	65.822	.562	66.384
2514	44.320	50.705	2.207	7.072	.669	104.973	.577	105.550
2515	50.962	63.792	2.239	8.511	.906	126.410	.811	127.221

Z_1 = เลขหมายโทรศัพท์ประเภทร้านค้า

Z_2 = เลขหมายโทรศัพท์ประเภทบ้านพัก

Z_3 = เลขหมายโทรศัพท์ประเภทพิเศษ

Z_4 = เลขหมายโทรศัพท์ประเภทราชการ

Z_5 = เลขหมายโทรศัพท์ประเภทสาธารณะ

Y'_1 = เลขหมายโทรศัพท์รวม 5 ประเภท

Z_6 = เลขหมายโทรศัพท์ประเภท ทศท.

Y_1 = เลขหมายโทรศัพท์รวมทุกประเภท

การหาเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่ารวมทุกประเภทนี้จะหาทั้งเส้นแนวโน้มที่เป็นเส้นตรง (Linear Trend) และเส้นแนวโน้มที่ไม่เป็นเส้นตรง (Non - linear Trend) ดังนี้

แสดงสมการต่างๆที่ใช้ในการประมาณ
จำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่ารวมทุกประเภท

สมการ	Standard error of estimate	เปอร์เซ็นต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Y}_1 = 20.718111 + 10.483250X$	15.895316	78.85
(2) $\hat{Y}_1 = 41.135506 - 7.017374X + 2.187578X^2$	6.967489	96.52
(3) $\hat{Y}_1 = (30.58)(1.167)^X$	15.124177	80.85
(4) $\hat{Y}_1 = 35.854702 + (1.592369)(1.674)^X$	5.323039	97.63
(5) $\hat{Y}_1 = (33.70)(1.098)^{1.403X}$	7.227808	96.25
(6) $\hat{Y}_1 = \frac{1}{33331.706006-(6055.930622)(1.204)^X}$	11.004267	89.86

จะเห็นว่าสมการที่(4) คือ $\hat{Y}_1=35.854702+(1.592369)(1.674)^X$ ซึ่งเป็น modified exponential curve ให้ค่า standard error of estimate น้อยที่สุดคือ 5.323039 และให้ค่า 97.63% ของการกระจายของตัวแปรไม่อิสระซึ่งมากที่สุด แสดงว่าเมื่อใช้สมการนี้ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่า จะให้ค่าประมาณ ($\hat{Y}_1$)ที่ใกล้เคียงกับค่าจริง (true value คือ Y_1) มากที่สุด แต่ปรากฏว่าเมื่อใช้สมการนี้พยากรณ์จำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่าในอนาคตแล้ว จะให้ค่าที่ มากเกินความเป็นจริง กล่าวคือ ค่าพยากรณ์ที่เพิ่มขึ้นในแต่ละปีจะเพิ่มขึ้นมากเป็นเท่าตัว ซึ่งตามข้อเท็จจริงแล้ว ผู้ต้องการเช่าโทรศัพท์ที่เพิ่มขึ้นในแต่ ละปีควรจะเป็นสัดส่วนกันมิใช่เพิ่มขึ้นทีละเท่าตัว ที่เป็นเช่นนี้เพราะ

อัตราการเพิ่มขึ้น (rate of increase) สูงคือ ค่า b เท่ากับ 1.674 จึงทำให้ค่าพยากรณ์ของจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่าเพิ่มขึ้นปีละเท่าตัว ดังนั้นเราจึงเลือกใช้สมการอื่นที่ให้ค่า standard error of estimate มากกว่าสมการที่ (4) แต่น้อยกว่าสมการอื่นๆ และค่าเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระที่น้อยกว่าสมการที่ (4) แต่มากกว่าสมการอื่นๆ กรณีนี้คือ สมการที่ (2)

$$\widehat{Y}_1 = 41.135506 - 7.017374X + 2.187578\, X^2$$

ซึ่งมี curve เป็น second - degree polynomial ให้ค่า standard error of estimate เท่ากับ 6.967489 และให้ค่า 96.52% ของการกระจายของตัวแปรไม่อิสระ เมื่อใช้สมการที่ (2) ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่าจะให้ค่าประมาณ ($\widehat{Y}_1$) ใกล้เคียงกับค่าจริง (Y_1) เช่นกันดังตารางที่ 2.1.2

ตารางที่ 2.1.2

แสดงการเปรียบเทียบค่าประมาณ ($\widehat{Y}_1$) กับค่าจริง (Y_1) เป็นรายปี

(พันเลขหมาย)

ปี	X	Y_1	$\widehat{Y}_1$
2507	0	35.064	41.136
2508	1	39.155	36.306
2509	2	42.060	35.851
2510	3	43.861	39.772
2511	4	48.170	48.067
2512	5	56.395	60.738
2513	6	66.384	77.784
2514	7	105.550	99.205
2515	8	127.221	125.002

ให้ X เป็น time ที่กำหนดขึ้น โดยกำหนดให้ X ของปี 2507 เท่ากับ 0 ปี 2508 เท่ากับ 1 ปี 2509 เท่ากับ 2 เช่นนี้ไปเรื่อยๆ

Y_1 เป็นค่าจริงของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท

$\hat{Y}_1$ เป็นค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท

ซึ่งประมาณจาก second-degree polynomial คือ

$$Y = a + bX + cX^2 + e$$

หาค่า a,b,c โดยวิธี least squares ซึ่งมี normal equation ตามโมเดล (4) ดังนี้

$$563.860 = 9a + 36b + 204c$$

$$2884.435 = 36a + 204b + 1296c$$

$$18486.657 = 204a + 1296b + 8772c$$

ซึ่งแก้สมการหาค่า a,b,c ได้ดังนี้

$$a = 41.135506$$

$$b = -7.017374$$

$$c = 2.187578$$

ดังนั้นสมการที่ใช้ในการประมาณคือ

$$\hat{Y}_1 = 41.135506 - 7.017374X + 2.187578X^2$$

ซึ่งให้ค่า $s = \sqrt{\frac{291.275398}{6}} = 6.967489$ และ 96.52% ของการกระจายของตัวแปรไม่อิสระ จากสมการที่ (2) ทดสอบสมมติฐาน (test hypothesis) เกี่ยวกับค่า b,c โดยใช้ตาราง analysis of variance ดังนี้

ตารางที่ 2.1.3

แสดง analysis of variance เพื่อทดสอบค่า b

Source of variation	Degree of Freedom	Sum of Squares	Mean Squares	F
Due to a	1	$(\Sigma Y)^2/n$	$\frac{(\Sigma Y)^2}{n}/1$	
Due to b/a	1	$b\ \Sigma xy$	$b\ \Sigma xy/1$	$\frac{b\ \Sigma xy}{\Sigma(Y-\hat{Y})^2/(n-2)}$
Residual	n-2	$\Sigma(Y-\hat{Y})^2$	$\Sigma(Y-\hat{Y})^2/n-2$	
Total	n	ΣY^2		

$$x = X - \bar{X}, \quad y = Y - \bar{Y}$$

การทดสอบค่า b จะทดสอบจากสมการเส้นตรงก่อน คือ

$$Y = 20.718111 + 10.483250\ X$$

$$H_o : b = 0$$

Sum of Squares due to a $= \frac{(\Sigma Y)^2}{n} = 35326.455511$

Sum of Squares due to b/a $= b\,\Sigma xy = 6593.911833$

Residual Sum of Squares $= \Sigma(Y-\hat{Y})^2 = 1768.627420$

Total Sum of Squares $= \Sigma Y^2 = 43688{,}994764$

$$\text{Mean Squares} = \frac{\text{Sum of Squares}}{\text{Degree of Freedom}}$$

$$F = \frac{\text{Mean Square due to b/a}}{\text{Residual Mean Squares}}$$

$$= \frac{6593.911833}{252.661060} = 26.097855$$

จากตาราง F-distribution ที่ 1 (n-2) deqree of freedom คือ

$F(1,7) = 3.59$ ที่ $\alpha = .10$

$F(1,7) = 5.59$ ที่ $\alpha = .05$

$F(1,7) = 8.07$ ที่ $\alpha = .025$

$F(1,7) = 12.25$ ที่ $\alpha = .010$

$F(1,7) = 16.24$ ที่ $\alpha = .005$

ค่า F ที่คำนวณได้ มากกว่า F จากตาราง F distribution ที่ 1, n-2 degree of freedom ทุกค่าของ α ดังนั้น reject H_0 ทุกค่าของ α

นั่นคือค่า $b \neq 0$ แสดงว่าค่า b significant พอที่จะอยู่ในสมการ

หลังจากทดสอบค่า b แล้ว ต่อไปจะทดสอบค่า c ว่า significant หรือไม่ โดยตั้งสมมติฐานว่า $H_0 : c = 0$ ถ้า accept H_0 แสดงว่า สมการของ second-degree polynomial ควรจะเหลือเพียงสมการเส้นตรง นั่นคือ การประมาณที่ใช้สมการเส้นตรงจะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงแล้ว ไม่จำเป็นต้องเพิ่ม X^2 เข้าไปในสมการ แต่ถ้า reject H_0 แสดงว่า สมการ second-degree polynomial เหมาะสมที่จะใช้ในการประมาณ

การทดสอบค่า c ทดสอบโดยตาราง analysis of variance เช่นกัน ดังตารางที่ 2.1.4

ตารางที่ 2.1.4

แสดง analysis of variance เพื่อทดสอบค่า c

Source of Variation	Degree of Freedom	Sum of Squares	Mean Squares
Regression on b,c	2	$\sum(\hat{Y} - \bar{\hat{Y}})^2$	$\sum(\hat{Y} - \bar{\hat{Y}})^2 /2$
Residuals	n - 3	$\sum(Y - \hat{Y})^2$	$\sum(Y - \hat{Y})^2 / (n-3)$
Total	n - 1	$\sum(Y - \bar{Y})^2$	
Regression on b	1	$b\sum xy$	$b\sum xy / 1$
Regression on c	1	$\sum(\hat{Y} - \bar{\hat{Y}})^2 - b\sum xy$	$\left\{\sum(\hat{Y} - \bar{\hat{Y}})^2 - b\sum xy\right\} / 1$

จากสมการ $\hat{Y}_1 = 41.135506 - 7.017374X + 2.187578X^2$

$H_0: c = 0$

Sum of Squares on b,c $= \sum(\hat{Y}-\bar{\hat{Y}})^2 = 8071.263853$

Resedual Sum of Squares $= \sum(Y-\hat{Y})^2 = 291,275398$

Rdsidual Mean Squares $= \dfrac{\sum(Y-\hat{Y})^2}{n-3} = 48.5459$

Total Sum of Squares $= \sum(Y-\bar{Y})^2 = 8362.539251$

Sum of Squares on b $= 6593.911833$

Sum of Squares on c $=$ SS on b,c $-$ SS on b

$= 8071.263853 - 6593.911833$

$= 1477.352020$

$$F = \frac{1477.352020}{48.5459} = 30.432066$$

จากตาราง F-distribution ที่ 1, (n-3) degree of freedom คือ

$F(1,6) = 3.78$ ที่ $\alpha = .10$

$F(1,6) = 5.99$ ที่ $\alpha = ,05$

$F(1,6) = 8.81$ ที่ $\alpha = .025$

$F(1,6) = 13.75$ ที่ $\alpha = .010$

$F(1,6) = 18.63$ ที่ α .005

ค่า F ที่คำนวนได้ มากกว่าค่า F จากตาราง F distribution ที่ 1,6 degree of freedom ทุกค่าของ α

$\therefore$ reject H_o ทุกค่าของ α

นั่นคือ $c \neq 0$ แสดงว่า ค่า c significant พอที่จะอยู่ในสมการ และแสดงว่าข้อมูลที่มีอยู่คือ จำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท มีแนวโน้มที่อยู่ในรูป second-degree polynomial ดังดูได้จากกราฟแสดงภาพที่ 1 ในภาคผนวก ก.

2) <u>การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า</u> ประมาณจากเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้าในอดีต ซึ่งมีสถิติที่รวบรวมไว้เป็นรายปีในตารางที่ 2.1.1

Z_1 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า

การหาเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า จะหาทั้งเส้นแนวโน้มที่เป็นเส้นตรง (Linear Trend) และเส้นแนวโน้มที่ไม่เป็นเส้นตรง (Non-Linear Trend) ดังต่อไปนี้ และพิจารณาดูว่าข้อมูลที่มีอยู่เหมาะสมกับ curve ใดในการประมาณและการพยากรณ์

แสดงสมการต่างๆที่ใช้ในการประมาณ
จำนวนเลขหมายโทรศัพท์ที่มีผู้เช่า ประเภทร้านค้า

สมการ	standard error of estimate	เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Z}_1 = 10.086232 + 4.263167\,X$	5.137872	85.51
(2) $\hat{Z}_1 = 15.52192 - 1.38194X + .70564\,X^2$	2.685567	96.61
(3) $\hat{Z}_1 = (15.86)(1.119)^X$	5.931328	80.69
(4) $\hat{Z}_1 = 13.6608 + (2.0640)(1.441)^X$	1.819156	98.09
(5) $\hat{Z}_1 = (11.62)(1.337)^{1.223^X}$	2.884718	95.44
(6) $\hat{Z}_1 = \frac{1}{140697.47505 - (75322.67758)(1.061)^X}$	2.700770	95.99

ปรากฏว่า สมการที่(4) คือ modified exponential curve ให้ค่า standard error of estimate น้อยที่สุดคือ 1.819156 และเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระสูงสุด คือ 98.09 แสดงว่าค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า ($\hat{Z}_1$) จาก modified exponential curve ใกล้เคียงกับค่าจริง (Z_1) มากกว่าใช้ curve อื่น แต่เมื่อใช้พยากรณ์จำนวนเลขหมายที่มีผู้ต้องการเช่าประเภทร้านค้าในอนาคตแล้ว เมื่อถึงระยะเวลาหนึ่งจะให้ค่าพยากรณ์เพิ่มขึ้นทีละเท่าตัว ดังนั้นด้วยเหตุผลเดียวกันกับการประมาณจำนวนเลขหมายที่มีผู้เช่ารวมทุกประเภท เราจะใช้สมการที่ (2) คือ second-degree polynomial ในการประมาณ ซึ่งให้ค่า standard error of estimate

เท่ากับ 2.685567 และ 96.61% ของการกระจายของตัวแปรไม่อิสระ โดยให้ค่าประมาณ ($\hat{Z}_1$) ใกล้เคียงกับค่าจริง (Z_1) ดังตารางที่ 2.1.5

$\hat{Z}_1$ จากตาราง 2.1.5 เป็น ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า ซึ่งประมาณจาก second-degree polynomial

$$Z_1 = a+bX+cX^2+e$$

แก้สมการหาค่า a,b,c โดยวิธี least squares เช่นเดียวกับการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท จะได้ดังนี้

$$a = 15.52192$$

$$b = -1.38194$$

$$c = .70564$$

ดังนั้น สมการที่ใช้ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า จะอยู่ในรูปดังนี้

$$\hat{Z}_1 = 15.52192 - 1.38194\ X + .70564\ X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b,c $H_o : b = 0$ โดยทำตาราง analysis of variance เหมือนตารางที่ 2.1.3 ได้ค่า F เท่ากับ 41.309443 ที่ 1,7 degree of freedom ซึ่งมากกว่า F ในตาราง F-distribution ที่ 1,7 degree of freedom ทุกค่าของ α เพราะฉะนั้น reject H_o นั่นคือ $b \neq 0$

ทดสอบสมมติฐานเกี่ยวกับค่า c $H_o : c=0$ ทำตาราง analysis of variance เหมือนตารางที่ 2.1.4 ได้ค่า F เท่ากับ 19.620801 ที่ 1,6 degree of freedom ซึ่งมากกว่า F ในตาราง F-distribution ที่ 1,6 degree of freedom ทุกค่าของ α ดังนั้น reject H_o แสดงว่า $c \neq 0$ นั่นคือ ค่า c significant พอที่จะอยู่ในสมการและข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า มีแนวโน้มในลักษณะของ second-degree polynomial ดังกราฟแสดงภาพที่ 2 ในภาคผนวก ก.

ตารางที่ 2.1.5

แสดงค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าแยกประเภทโดย Trend Method

(พันเลขหมาย)

ปี	X	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_3$	$\hat{Z}_4$	$\hat{Z}_5$	$\hat{Y}'_1$
2507	0	15.522	19.014	0.887	3.623	0.282	40.810
2508	1	14.846	14.616	1.004	3.880	0.262	35.978
2509	2	15.581	13.000	1.137	4.245	0.269	35.507
2510	3	17.727	14.166	1.287	4.719	0.303	39.397
2511	4	21.284	18.115	1.458	5.300	0.362	47.649
2512	5	26.253	24.847	1.651	5.990	0.447	60.261
2513	6	32.633	34.361	1.869	6.789	0.558	77.235
2514	7	40.425	46.657	2.117	7.695	0.696	98.570
2515	8	49.627	61.735	2.397	8.710	0.860	124.266

X กำหนดขึ้นตามปีโดยให้เท่ากับ 0,1,2,...,8 ตามลำดับ

$\hat{Z}_1$ = ค่าประมาณของเลขหมายโทรศัพท์ประเภทร้านค้า

$\hat{Z}_2$ = ค่าประมาณของเลขหมายโทรศัพท์ประเภทบ้านพัก

$\hat{Z}_3$ = ค่าประมาณของเลขหมายโทรศัพท์ประเภทพิเศษ

$\hat{Z}_4$ = ค่าประมาณของเลขหมายโทรศัพท์ประเภทราชการ

$\hat{Z}_5$ = ค่าประมาณของเลขหมายโทรศัพท์ประเภทสาธารณะ

$\hat{Y}'_1$ = ค่าประมาณของเลขหมายโทรศัพท์รวม 5 ประเภท

3) การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่าประเภทบ้านพัก ประมาณจากเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพัก ซึ่งมีสถิติรวบรวมไว้เป็นรายปีดังตารางที่ 2.1.1.

Z_2 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพัก

การหาเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพัก จะหาทุกสมการคือ สมการที่เป็นเส้นตรงและไม่เป็นเส้นตรง ดังนี้

แสดงสมการต่าง ๆ ที่คำนวณเพื่อใช้ประมาณ
จำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพัก

สมการ	standard error of estimate	เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Z}_2 = 6.029332 + 5.340167X$	9.297608	75.43
(2) $\hat{Z}_2 = 19.01376 - 5.78935X + 1.39119X^2$	5.081906	93.71
(3) $\hat{Z}_2 = (12.01)(1.187)^X$	8.808901	77.95
(4) $\hat{Z}_2 = 15.7012 + (.1647)(2.064)^X$	4.188165	95.02
(5) $\hat{Z}_2 = (15.31)(1.027)^{1.672^X}$	6.446075	88.19
(6) $\hat{Z}_2 = \dfrac{1}{68111.26916 - (3957.6778)(1.401)^X}$	16.914561	18.67

สมการที่ (4) คือ modified exponential curve เหมาะสมที่ใช้ในการประมาณ เพราะให้ค่า standard error of estimate น้อยที่สุด และเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระมากที่สุด ซึ่งถูกต้องด้วยเหตุผลทางทฤษฎี แต่เมื่อใช้หาค่าพยากรณ์ในอนาคตแล้ว จะให้ค่ามากเกินความเป็นจริง เพราะ b = 2.064 ดังนั้นเราจึงเลือกสมการอื่นในการประมาณ ตามเหตุผลเดียวกันกับการเลือกสมการที่

ใช้ในการประมาณ y_1 คือสมการที่ให้ค่า standard error of estimate น้อยที่สุดในทุก ๆ สมการยกเว้นสมการที่ (4) และค่าเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระมากที่สุดในทุก ๆ สมการยกเว้นสมการที่ (4) เช่นกัน สมการดังกล่าวนี้ได้แก่สมการที่ (2)

$$\hat{Z}_2 = 19.01376 - 5.78935X + 1.39119X^2$$

ซึ่งเป็น second-degree polynomial ให้ค่า standard error of estimate เท่ากับ 5.081906 และ 93.71 % ของการกระจายของตัวแปรไม่อิสระ โดยจะให้ค่าพยากรณ์ใกล้เคียงกับความเป็นจริงและค่าประมาณใกล้เคียงกับค่าจริงด้วย ตามตารางที่ 2.1.5

$\hat{Z}_2$ จากตารางที่ 2.1.5 เป็นค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพักที่ประมาณจากสมการ second-degree polynomial

$$Z_2 = a+bX+cX^2+e$$

โดยวิธี least squares เหมือนกับการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท solve หาค่า a,b,c ได้ดังนี้

$$a = 19.01376$$

$$b = -5.78935$$

$$c = 1.39119$$

ดังนั้น สมการจะอยู่ในรูปดังนี้

$$\hat{Z}_2 = 19.01376 - 5.78935X + 1.39119X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า significant ของ b โดยทำตาราง analysis of variance เช่นเดียวกับตารางที่ 2.1.3 ได้ค่า F =21.484272 ที่ (1,7) degree of freedom ซึ่งมากกว่า F (1,7) จากตาราง F-distribution ทุกค่า α ดังนั้น reject H_o: b = 0 แสดงว่า ค่า **increment** ของ Z_2 significant

ทดสอบสมมติฐานเกี่ยวกับค่า significant ของ c ว่า ข้อมูลที่มีอยู่มีแนวโน้มเป็น second-degree polynomial หรือไม่ โดยทำตาราง analysis of variance เหมือนตารางที่ 2.1.4 ได้ค่า F = 17.430809 ที่ (1,6) degree of freedom ซึ่งมากกว่า F (1,6) จากตาราง F - distribution ที่ α = .10, .05, .025, .010 นั่นคือ reject H_0 c = 0 หมายความว่าค่า c significant ดังนั้น ข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพักมีแนวโน้มเป็น second-degree polynomial ซึ่งจะดูภาพประกอบด้วยได้ดังกราฟแสดงภาพที่ 3 ในภาคผนวก ก.

4) การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่าประเภทพิเศษ ประมาณจากเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ โดยมีสถิติรวบรวมไว้เป็นรายปีตามตารางที่ 2.1.1

Z_3 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ

การหาเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ จะหาทุกสมการ ทั้งที่เป็นเส้นตรงและไม่เป็นเส้นตรง

แสดงสมการต่าง ๆ ที่คำนวณเพื่อใช้
ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ

สมการ	Standard error of estimate	เปอร์เซ็นต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Z}_3 = .779067 + .19215X$	.257876	82.64
(2) $\hat{Z}_3 = .898612 + .089683X + .012808X^2$	.262876	84.53
(3) $\hat{Z}_3 = (.8867)(1.132)^X$	.247760	83.97
(4) $\hat{Z}_3 = .962502 + (.051984)(1.573)^X$	.363863	65.43
(5) $\hat{Z}_3 = (.9091)(1.102)^{1.373^X}$	.416876	54.62
(6) $\hat{Z}_3 = \frac{1}{1215912.4026-(206596.9429)(1.207)^X}$	.547771	46.53

จะเห็นว่า สมการที่ (3)

$$\hat{Z}_3 = (.8867)(1.132)^X$$

ซึ่งเป็น simple exponential curve ให้ค่า standard error of estimate ต่ำสุด คือ .247760 และเปอร์เซ็นต์ของการกระจายของตัวแปรไม่อิสระสูงสุด คือ 83.97 % และเมื่อใช้สมการดังกล่าวคำนวณหาค่าพยากรณ์ของจำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่าในอนาคต ปรากฏว่าใกล้เคียงกับความเป็นจริง สมการที่ (3) นี้จึงเป็นสมการที่เหมาะสมที่สุดในการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ ดังนั้นเราจะใช้สมการที่ (3) นี้ในการประมาณเพราะค่าประมาณ ($\hat{Z}_3$) ใกล้เคียงกับค่าจริง (Z_3) ด้วย ดังตารางที่ 2.1.5

$\hat{Z}_3$ จากตาราง 2.1.5 เป็นค่าประมาณของ Z_3 ที่ประมาณจาก simple exponential curve

$$Z_3 = ab^X$$

การ solve หาค่า a,b ต้องแปลงสมการให้อยู่ในรูปเส้นตรง โดย take logarithm ทั้ง 2 ข้าง

$$\log Z_3 = \log a + X \log b \quad *$$

โดยวิธี least equares ได้ normal equation ตามโมเดล 6 หน้า 11 ดังนี้

$$1.4739 = 9 \log a + 36 \log b$$

$$9.1355 = 36 \log a + 204 \log b$$

solve หาค่า log a, log b ได้ดังนี้

$$\log a = \bar{1}.947775$$

$$\log b = .053998$$

$$\therefore \log \hat{Z}_3 = \bar{1}.947775 + .053998\, X$$

เปิดตาราง anti-logarithm ได้ค่า Z_3, a, b

$$\hat{Z}_3 = (.8867)(1.132)^X$$

ทดสอบสมมติฐานเกี่ยวกับค่า log b โดยตาราง analysis of variance ซึ่งแสดงในตารางที่ 2.1.6 ดังนี้

ตารางที่ 2.1.6

แสดง analysis of variance เพื่อทดสอบค่า log b

Source of variation	degree of freedom	Sum of Squares	Mean Squares	F
Due to a'	1	$(\Sigma Z_3')^2/n$	$\frac{(\Sigma Z_3')^2}{n}/1$	
Due to b'/a'	1	$b'\Sigma xz_3'$	$b'\Sigma xz_3'/1$	$\frac{b'\Sigma xz_3'/1}{\Sigma(Z_3'-\hat{Z}_3')^2/(n-2)}$
Residual	n-2	$\Sigma(Z_3'-\hat{Z}_3')^2$	$\Sigma(Z_3'-\hat{Z}_3')^2/(n-2)$	
Total	n	$\Sigma Z_3'^2$		

จาก * ให้

$$\log Z_3 = Z_3'$$

$$\log \hat{Z}_3 = \hat{Z}_3'$$

$$\frac{\Sigma \log Z_3}{N} = \bar{Z}_3'$$

$$Z_3'-\bar{Z}_3' = z_3'$$

$$\log a = a'$$

$$\log b = b'$$

Sum of Squares due to a' = .241376

Sum of Squares due to b'/a' = .174949

Residual Sum of Squares = .022724

Total Sum of Squares = .439049

H_o: log b = 0

$$F = \frac{\text{Mean Square due to } b'/a'}{\text{Residual Mean Squares}}$$

$$= \frac{.174949}{.022724/7} = 53.896796$$

ค่า F ที่คำนวณได้นี้มากกว่า F (1,7) ในตาราง F-distribution ทุกค่าของ $\propto$ นั่นคือเราจะ reject H_0 : log b = 0 แสดงว่า log b significant ในสมการ

$$\log \hat{Z}_3 = \bar{4}.947775 + .053998\ X$$

ดังนั้น ในการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ จะใช้สมการ (3) คือ simple exponential จะให้ค่าประมาณใกล้เคียงกับค่าจริงที่สุด ดูกราฟแสดงภาพที่ 4 ประกอบด้วย (ในภาคผนวก ก.)

5) <u>การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการ</u> ประมาณจากข้อมูลจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการ ที่รวบรวมไว้จากตารางที่ 2.1.1

Z_4 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการ

วิธีการประมาณคือ หาเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการว่าอยู่ในลักษณะใด ดังนั้นจึงหาเส้นแนวโน้มทั้งที่เป็นเส้นตรงและไม่เป็นเส้นตรง แล้วพิจารณาดูว่าเส้นแนวโน้มเส้นใดเหมาะสม (fit) กับข้อมูล

แสดงสมการต่าง ๆ ที่คำนวณเพื่อใช้ในการประมาณ
จำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการ

สมการ	standard error of estimate	เปอร์เซ็นต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Z}_4 = 3.1176 + .635933\ X$	.566384	89.32
(2) $\hat{Z}_4 = 3.622983 + .202748\ X + .054148X^2$	.449934	94.22
(3) $\hat{Z}_4 = (3.693)(1.104)^X$	.458657	94.16
(4) $\hat{Z}_4 = 3.736829 + (1.344409)(1.406)^X$	.491413	93.11
(5) $\hat{Z}_4 = (3.580)(1.134)^X$	.499570	91.69
(6) $\hat{Z}_4 = \frac{1}{299328.7556-(52203.339887)(1.18)^X}$	.658643	85.55

วิธีการตัดสินใจว่าจะใช้สมการใดในการประมาณ คงใช้หลักการเดียวกับการตัดสินใจเลือกใช้สมการในการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทต่าง ๆ ที่กล่าวมาแล้ว คือสมการที่ให้ค่า standard error of estimate ต่ำสุด และเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระสูงสุด จะเห็นว่าสมการที่ (2) second-degree polynomial เป็นสมการที่อยู่ในหลักเกณฑ์ เพราะค่า standard error of estimate เท่ากับ .449934 ซึ่งน้อยที่สุดและ 94.22 % ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ดังนั้นสมการที่ (2) นี้จะให้ค่าประมาณที่ใกล้เคียงกับค่า**จริง**ที่สุด ซึ่งแสดงไว้ในตารางที่ 2.1.5

$\hat{Z}_4$ ในตารางที่ 2.1.5 เป็นค่า**ประ**มาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการซึ่งคำนวณจาก second-degree polynomial

$$Z_4 = a + bX + cX^2 + e$$

โดยวิธี least squares หาค่า a,b,c ได้ดังนี้

$$a = 3.622983$$

$$b = .202748$$

$$c = .054148$$

ดังนั้นสมการที่ใช้ประมาณคือ

$$\hat{Z}_4 = 3.622983 + .202748X + .054148X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b และ c โดยทำ analysis of variance ตามที่แสดงไว้ในตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_o: b = 0$

$F = 58.5281$ มี (1,7) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่าค่า F (1,7) จากตาราง F-distribution ทุกค่าของ α ดังนั้น reject $H_o: b = 0$ นั่นคือ $b \neq 0$

$H_o: c = 0$

$F = 5.092306$ มี (1,6) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่าค่า F (1,6) จากตาราง F - distribution ที่ α= .10 ดังนั้น reject H_o: c = 0 นั่นคือ c $\neq$ 0 แสดงว่าสมการ second-degree polynomial ใช้ได้กับการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการ หรืออีกนัยหนึ่งแสดงว่าข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการมีเส้นแนวโน้ม (Trend line) เป็น second-degree polynomial ดูกราฟแสดงภาพที่ 5 ในภาคผนวก ก.

6) <u>การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะ</u> ประมาณจากข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะที่แสดงไว้ในตารางที่ 2.1.1

Z_5 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะ
วิธีการประมาณคือ นำข้อมูล Z_5 จากตารางที่ 2.1.1 มาหาเส้นแนวโน้มทั้งที่เป็นเส้นตรงและไม่เป็นเส้นตรง แล้วพิจารณาดูว่าสมการใดเหมาะสมกับข้อมูล

แสดงสมการต่าง ๆ ที่ใช้คำนวณหาเส้นแนวโน้ม
ของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะ

สมการ	Standard error of estimate	เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Z}_5 = .159778 + .07225X$	.093712	83.60
(2) $\hat{Z}_5 = .281535 - .032113X + .013045X^2$	.038833	97.59
(3) $\hat{Z}_5 = (.2226)(1.165)^X$	.065765	91.92
(4) $\hat{Z}_5 = .222 + (.034697)(1.442)^X$	.049621	97.28
(5) $\hat{Z}_5 = (.1866)(1.361)^{1.224^X}$	.041663	97.40
(6) $\hat{Z}_5 = \dfrac{1}{9424927.1752-(5447152.5567)(1.054)^X}$	.108157	78.15

การพิจารณาเลือกว่าจะใช้สมการใดนั้น ก็ใช้หลักการเดียวกับที่กล่าวมาแล้ว ในการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท สมการที่ (2) second-degree polynomial อยู่ในหลักเกณฑ์ที่จะนำมาใช้ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะ เพราะให้ค่า standard error of estimate เท่ากับ .038833 ซึ่งต่ำสุด และ 97.59 เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ดังนั้น สมการที่ (2) จะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงที่สุด

โดย $\hat{Z}_5$ จากตาราง 2.1.5 เป็นค่าประมาณของ Z_5 ซึ่งคำนวณจาก second-degree polynomial

$$Z_5 = a + bX + cX^2 + e$$

โดยวิธี least squares แก้สมการหาค่า a,b,c ได้ดังนี้

$$a = .281535$$

$$b = -.032113$$

$$c = .013045$$

ดังนั้น สมการ second-degree polynomial ที่ใช้ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะคือ

$$\hat{Z}_5 = .281535 - .032113X + .013045X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b,c โดยทำ analysis of variance ที่แสดงไว้ในตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_0 : b = 0$

$F = 35.663238$ มี (1,7) degree of freedom

F ที่คำนวณได้นี้มากกว่า F จากตาราง F-distribution ที่ (1,7) of freedom ทุกค่าของ α ดังนั้นเรา reject H_0: $b = 0$ นั่นคือ $b \neq 0$

$H_0 : c = 0$

$F = 34.755589$ มี (1,6) degree of freedom

F ที่คำนวณได้นี้มากกว่า F จากตาราง F-distribution ที่ (1,6) degree of freedom ทุกค่าของ α ดังนั้นเรา reject H_o: c = 0 นั่นคือ $c \neq 0$ แสดงว่าสมการ second-degree polynomial เหมาะสมที่จะใช้ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะ หรืออีกนัยหนึ่งคือ ข้อมูลชนิดนี้มีแนวโน้มที่อยู่ในลักษณะ second-degree polynomial ดูกราฟแสดงภาพที่ 6 ภาคผนวก ก

7) <u>การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท.</u> เนื่องจากข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท.นี้ fluctuate มาก ดังที่แสดงไว้ในตารางที่ 2.1.1 เช่น ในปี 2509 มี 397 เลขหมาย ปี 2510 เพิ่มขึ้นเป็น 573 เลขหมาย แต่ในปี 2511 ลดลงเป็น 343 เลขหมาย ต่อมาปี 2512 เพิ่มขึ้นเพียง 5 เลขหมาย เป็น 348 เลขหมาย เป็นต้น ดังนั้นถ้าใช้ข้อมูลประเภททศท.คำนวณหาเส้นแนวโน้มเพื่อใช้ในการประมาณแล้ว จะไม่ให้ค่าประมาณที่ดี เพราะฉะนั้นจะใช้วิธีประมาณจากจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวม 5 ประเภท คือ ร้านค้า+ บ้านพัก + พิเศษ+ ราชการ + สาธารณะ แล้วนำไปหักออกจากจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท คือ ร้านค้า + บ้านพัก + พิเศษ + ราชการ + สาธารณะ + ทศท. ก็จะได้จำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท. ข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. แสดงไว้ในตารางที่ 2.1.1

Y_1' เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท.

วิธีประมาณ คือหาเส้นแนวโน้มของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. จากสมการต่าง ๆ ทั้งที่เป็นเส้นตรงและไม่เป็นเส้นตรง แล้วพิจารณาดูว่าสมการใดเหมาะสมและให้เหตุผลที่ดีกับข้อมูล

แสดงสมการต่าง ๆ ที่ใช้คำนวณหาค่าประมาณ
ของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท.

สมการ	Standard error of estimate	เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Y}'_1 = 20.458 + 10.432X$	15.826006	78.34
(2) $\hat{Y}'_1 = 40.81015-7.0127X + 2.1805876X^2$	6.967729	96.48
(3) $\hat{Y}'_1 = (30.27)(1.168)^X$	12.246947	87.33
(4) $\hat{Y}'_1 = 35.57367 + (1.564399)(1.676)^X$	5.332208	97.60
(5) $\hat{Y}'_1 = (33.50)(1.096)^{1.406^X}$	7.607290	96.31
(6) $\hat{Y}'_1 = \frac{1}{33478.37705-(5969.31213)(1.206)^X}$	10.156114	91.29

การพิจารณาเลือกใช้สมการใดในการประมาณ คงใช้หลักการเดียวกันกับที่กล่าวมาแล้วในตอนที่ (1) ของข้อ 2.1.1 สมการที่ (4) modified exponential curve ให้ค่า standard error of estimate เท่ากับ 5.332208 ซึ่งต่ำสุด และ 97.60 เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด แต่ด้วยเหตุผลเดียวกันกับในตอนที่ (1) ของข้อ 2.1.1 คือ ถ้าใช้สมการที่ (4) ประมาณจะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงก็ตาม แต่เมื่อใช้พยากรณ์จำนวนเลขหมายโทรศัพท์ที่มีผู้ต้องการเช่าในอนาคต ค่าพยากรณ์ที่ได้จะไม่สมเหตุผล ดังนั้น จึงเลือกใช้สมการอื่นที่ให้ค่า standard error of estimate น้อยที่สุด และเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระสูงสุด ในทุก ๆ สมการยกเว้นสมการที่ (4) ได้แก่ สมการที่ (2) second - degree polynomial ให้ค่า standard error of estimate เท่ากับ 6.967729 และ 96.48 เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระ ซึ่งจะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงมากกว่าค่าประมาณจากสมการอื่นยกเว้นสมการที่ (4) ตามตารางที่ 2.1.5

$\hat{Y}_1'$ จากตาราง 2.1.5 เป็นค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. โดยคำนวณจากสมการ second - degree polynomial

$$Y_1' = a + bx + cx^2 + e$$

โดยวิธี least squares แก้สมการหาค่า a,b,c ได้ดังนี้

$$a = 40.81015$$

$$b = -7.0127$$

$$c = 2.1805876$$

ดังนั้น สมการ second - degree polynomialที่ใช้ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. คือ

$$\hat{Y}_1' = 40.81015-7.0127X + 2.1805876X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b,c โดยทำ analysis of variance ที่แสดงไว้ในตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_o : b = 0$

$F = 26.082154$ มี (1,7) degree of freedom

ค่า F มากกว่า F (1,7) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject $H_o : b = 0$ นั่นคือ $b \neq 0$

$H_o : c = 0$

$F = 30.112542$ มี (1,6) degree of freedom

ค่า F มากกว่า F (1,6) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject $H_o : c = 0$ นั่นคือ $c \neq 0$ แสดงว่า สมการที่ (2) นี้เหมาะสมที่จะใช้ประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. ดูกราฟแสดงภาพที่ 7 ประกอบด้วย (ในภาคผนวก ก)

เมื่อได้ค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท.แล้วก็นำไปหักจากค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท จะได้ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท. ตามตารางที่ 2.1.7 ดังนี้

ตารางที่ 2.1.7

แสดงค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท.

(พันเลขหมาย)

ปี	$\hat{Y}_1$	$\hat{Y}'_1$	$\hat{Y}_1-\hat{Y}'_1$
2507	41.136	40.810	.326
2508	36.306	35.978	.328
2509	35.851	35.507	.344
2510	39.772	39.397	.375
2511	48.067	47.649	.418
2512	60.738	60.261	.477
2513	77.784	77.235	.549
2514	99.205	98.570	.635
2515	125.002	124.266	.736

$\hat{Y}_1$ ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท(จากตารางที่ 2.1.2)

$\hat{Y}'_1$ ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. (จากตารางที่ 2.1.5)

$\hat{Y}_1 - \hat{Y}'_1$ ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท.

วิธีการปรับตัวเลขโดย correction factors

เนื่องจาก การประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท. คำนวณจากวิธีดังกล่าวในข้อ 2.1.7 ทำให้ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท ไม่เท่ากับผลบวกของค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่า

แต่ละประเภท ดังนั้น จึงต้องปรับตัวเลขค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าแต่ละประเภท เพื่อให้ได้ผลรวมเท่ากับ ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท โดยวิธีทำ correction factor

การปรับตัวเลขค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าแต่ละประเภทนั้น เราจะปรับเพียง 5 ประเภท คือ ประเภทร้านค้า, บ้านพัก, พิเศษ, ราชการ, สาธารณะ สำหรับประเภททศท.เราจะไม่ปรับ แต่จะใช้ตัวเลขเดิมที่ประมาณได้จากตารางที่ 2.1.7 เพราะเป็นค่าประมาณที่ประมาณได้จากผลต่างระหว่างค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท กับค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. ดังนั้นเราจะคำนวณ correction factor ระหว่างผลรวมของค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่า 5 ประเภท กับค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. ดังตารางที่ 2.1.8 นี้

ความหมายของ column ในตารางที่ 2.1.8 คือ

$\hat{Z}_1$ = ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า (ได้จากตารางที่ 2.1.5)

$\hat{Z}_2$ = ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพัก (ได้จากตารางที่ 2.1.5)

$\hat{Z}_3$ = ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ (ได้จากตารางที่ 2.1.5)

$\hat{Z}_4$ = ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการ (ได้จากตารางที่ 2.1.5)

$\hat{Z}_5$ = ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะ (ได้จากตารางที่ 2.1.5)

$Y''_1 = \hat{Z}_1 + \hat{Z}_2 + \hat{Z}_3 + \hat{Z}_4 + \hat{Z}_5$

Y'_1 = ค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมยกเว้นทศท. (ได้จากตารางที่ 2.1.5)

C = correction factors ได้จาก Y'_1/Y''_1

ตารางที่ 2.1.8

แสดง correction factors ของค่าประมาณเลขหมายโทรศัพท์ที่มีผู้เช่า

(พันเลขหมาย)

ปี	$\hat{Z}_1$	$\hat{Z}_2$	$\hat{Z}_3$	$\hat{Z}_4$	$\hat{Z}_5$	$\hat{Y}''_1$	$\hat{Y}'_1$	C
2507	15.522	19.014	.887	3.623	.282	39.328	40.810	1.03768
2508	14.846	14.616	1.004	3.880	.262	34.608	35.978	1.03958
2509	15.581	13.000	1.137	4.245	.269	34.232	35.507	1.037246
2510	17.727	14.166	1.287	4.719	.303	38.202	39.397	1.03128
2511	21.284	18.115	1.458	5.300	.362	46.519	47.649	1.02429
2512	26.253	24.847	1.651	5.990	.447	59.188	60.261	1.018129
2513	32.633	34.361	1.869	6.789	.558	76.210	77.235	1.01344
2514	40.425	46.657	2.117	7.695	.696	97.590	98.570	1.01004
2515	49.627	61.735	2.397	8.710	.860	123.329	124.266	1.00759

เมื่อได้ค่า correction factors แล้วก็นำมาปรับค่าประมาณของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า, บ้านพัก, พิเศษ, ราชการ, สาธารณะ โดยนำค่า แต่ละปีคูณกับค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าแต่ละประเภทในตารางที่ 2.1.8 เป็นรายปี จะให้ค่าประมาณใหม่ตามตารางที่ 2.1.9 ดังนั้นผลรวมของค่าประมาณใหม่ของจำนวนเลขหมายโทรศัพท์ 5 ประเภทกับค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท. จะเท่ากับค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท

ความหมายของ column ต่าง ๆ ในตารางที่ 2.1.9 คือ

$\hat{Z}'_1$ = ค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทร้านค้า โดยวิธี correction factor

$\hat{Z}'_2$ = ค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทบ้านพัก โดยวิธี correction factor

$\hat{Z}'_3$ = ค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทพิเศษ โดยวิธี correction factor

$\hat{Z}'_4$ = ค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทราชการ โดยวิธี correction factor

$\hat{Z}'_5$ = ค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภทสาธารณะ โดยวิธี correction factor

$\hat{Z}_6$ = ค่าประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท.

$Y_1 = \hat{Z}'_1 + \hat{Z}'_2 + \hat{Z}'_3 + \hat{Z}'_4 + \hat{Z}'_5 + \hat{Z}_6$ ซึ่งเท่ากับค่า Y_1 ในตารางที่ 2.1.2

ตารางที่ 2.1.9

แสดงค่าประมาณใหม่ของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่า 5 ประเภท โดยวิธี **correction factor**

(พันเลขหมาย)

ปี	$\hat{Z}'_1$	$\hat{Z}'_2$	$\hat{Z}'_3$	$\hat{Z}'_4$	$\hat{Z}'_5$	$\hat{Z}_6$	$\hat{Y}_1$
2507	16.107	19.731	.920	3.759	.293	.326	41.136
2508	15.434	15.195	1.044	4.033	.272	.328	36.306
2509	16.162	13.484	1.179	4.403	.279	.344	35.851
2510	18.282	14.609	1.327	4.867	.312	.375	39.772
2511	21.801	18.555	1.493	5.429	.371	.418	48.067
2512	26.729	25.297	1.681	6.099	.455	.477	60.738
2513	33.072	34.823	1.894	6.880	.566	.549	77.784
2514	40.831	47.126	2.138	7.772	.703	.635	99.205
2515	50.004	62.204	2.415	8.776	.867	.736	125.002

2.1.2 วิธีการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าโดย Multiple Linear Regression Method

เป็นการประมาณโดยใช้ปัจจัยต่าง ๆ ที่เกี่ยวข้องกับเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทในเขตนครหลวง มาสร้างเป็น regression function ในรูปกำลังหนึ่ง โดยใช้หลักการของ Step - wise multiple regression ตัวแปรต่าง ๆ ที่นำมาใช้ในการสร้างโมเดล มีดังต่อไปนี้

1. ตัวแปรที่ไม่อิสระ - Y_1 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทในเขตนครหลวง ซึ่งแสดงไว้ในตารางที่ 2.1.1
2. ตัวแปรอิสระ ได้แก่

   X_1 เป็นจำนวนประชากรในเขตนครหลวง

   X_2 เป็นจำนวนบ้านในเขตนครหลวง

   X_3 เป็นจำนวนธุรกิจในเขตนครหลวง

   X_4 เป็น G.D.P. ของประชากรในเขตนครหลวง

   X_5 เป็นความหนาแน่นของการใช้โทรศัพท์ ซึ่งหมายถึงจำนวนเครื่องโทรศัพท์ต่อประชากร 1,000 คน

ค่าของตัวแปรอิสระ $X_1, X_2 \ldots, X_5$ แสดงไว้ในตารางที่ 2.1.10

ตารางที่ 2.1.10

แสดงค่าของตัวแปรอิสระ X_1, X_2, X_3, X_4, X_5 เป็นรายปี

ปี	X_1 (หมื่น)	X_2 (พัน)	X_3 (พัน)	X_4 (พันล้าน)	X_5 (เครื่องต่อ 1,000 คน)
2507	27.41118	34.4143	65.258	20.9525	19.5854
2508	28.96352	36.6110	70.985	22.8524	19.8435
2509	30.01706	38.9479	76.889	25.5871	21.2162
2510	31.23602	41.4339	83.945	28.9884	23.9524
2511	32.47339	44.1311	91.061	32.2883	26.9211
2512	36.46878	50.7959	97.859	37.1255	28.5142
2513	37.89205	53.2470	103.909	40.8381	31.3545
2514	39.45254	55.6784	110.020	44.9219	40.8571
2515	40.96580	58.7233	117.716	49.4141	46.4133

ที่มาของข้อมูล

X_1 จากกองทะเบียนราษฎร์กระทรวงมหาดไทย

X_2 จากกองทะเบียนราษฎร์กระทรวงมหาดไทย

X_3 จากกรมทะเบียนการค้ากระทรวงพาณิชย์

X_4 จากวิทยานิพนธ์ ของ Somluckrat Wattanavitukul เรื่อง Effect of Income Equalization on Overall Economic Growth of Thailand P.44 แต่มีเพียงปี 2511 จึงคาดคะเนโดยประมาณด้วยอัตราการเพิ่มขึ้น เฉลี่ย 10 % ต่อปี ตั้งแต่ปี 2512 จนถึงปี 2515

วิธีการสร้างโมเดลโดยใช้หลักการของ Step-wise multiple regression จะให้เครื่องคอมพิวเตอร์ IBM SYSTEM 30 MODEL 40 คำนวณ โดยเครื่องจะคำนวณเป็นขั้น ๆ ดังนี้

ขั้นที่ 1 หาค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) ระหว่างจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทกับตัวแปรอิสระแต่ละตัว จะได้ correlation matrix ดังนี้

	Y_1	X_1	X_2	X_3	X_4	X_5
Y_1	1.00000	.89426	.88706	.88958	.91837	.98570
X_1	.89426	1.0000	.99926	.98948	.99402	.94163
X_2	.88706	.99926	1.00000	.99265	.99514	.93948
X_3	.88958	.98948	.99265	1.00000	.99590	.94683
X_4	.91837	.99402	.99514	.99590	1.00000	.96517
X_5	.98570	.94163	.93948	.94683	.96517	1.00000

จาก correlation matrix ตัวแปรอิสระที่มีค่าสัมประสิทธิ์สหสัมพันธ์สูงที่สุด คือ X_5 - ความหนาแน่นการใช้โทรศัพท์ เมื่อยกกำลังสองของค่าสัมประสิทธิ์สหสัมพันธ์ตัวที่สูงที่สุดก็จะให้ค่ากำลังสองสูงสุดด้วย ดังนั้น เครื่องคอมพิวเตอร์จะนำตัวแปรอิสระ X_5 มาสร้างโมเดล simple linear regression

$$Y_1 = -34.44496 + 3.37846\, X_5$$

$$\therefore\ a = -34.44496$$

$$b_1 = 3.37846$$

F - value for analysis of variance = 239.471

Standard error of estimate = 5.825

<u>ขั้นที่ 2</u> เครื่องคอมพิวเตอร์จะหาตัวแปรอิสระตัวต่อไปเพื่อนำมาใส่ในโมเดล โดยจะเลือกตัวแปรอิสระที่ให้ค่า V_2 สูงสุด คือ X_3 - จำนวนธุรกิจในเขตนครหลวง

$$V_2 = (1-r^2_{YX_5})r^2_{YX_3 \cdot X_5}$$

นั่นคือ ในขั้นที่ 1 เมื่อนำ X_5 มา explained Y_1 แล้ว สัดส่วนของกระจายของ Y_1 กับ X_3 เมื่อ X_5 อยู่ในสมการแล้วสูงสุด ดังนั้นเมื่อนำ X_3 เข้าในโมเดลเป็นตัวที่สอง จะได้โมเดลเป็น multiple linear regression คือ

$$Y_1 = -5.04971 + 4.74825X_5 - .75697X_3$$

$$a' = -5.04971$$

$$b'_1 = 4.74852$$

$$b_2 = .75697$$

F - value for analysis of variance = 298.494

Standard error of estimate = 3.724

เมื่อเปรียบเทียบค่า standard error of estimate ในขั้นที่ 1 กับขั้นที่ 2 แล้ว จะเห็นว่าเมื่อเรานำ X_3 เข้ามาช่วย explained ค่า Y_1 แล้ว ทำให้ค่า standard error of estimate ลดลง และค่า F ที่เครื่องคำนวณได้นี้จะเป็นการทดสอบ $H_0 : b_1 = b_2 = 0$ หมายความว่า ทดสอบรวม (over-all test) ว่าสมการ regression นี้ significant หรือไม่ ถ้าค่า F ยิ่งมากแสดงว่าสมการ regression นี้ significant คือผลต่างระหว่างค่าจริง Y กับค่าประมาณ Y น้อยมาก [5] ดังนั้นในขั้นที่ 2 นี้ค่า F = 298.494 มากกว่าค่า F ในขั้นที่ 1 แสดงว่าสมการ regression ในขั้นที่ 2 นี้จะให้ค่าประมาณ Y_1 นี้แตกต่างจากค่าจริง Y_1 น้อยกว่า สมการในขั้นที่ 1

<u>ขั้นที่ 3</u> ตัวแปรอิสระ X_1 - จำนวนประชากรในเขตนครหลวง เป็นตัวแปรที่ 3 ที่เครื่องคอมพิวเตอร์เลือกเข้ามาโดยใช้หลักการเลือกเช่นเดียวกับในขั้นที่ 2 ดังนั้นจะได้โมเดลดังนี้

---

[5] Taro Yamane, <u>Statistics; An Introductory Analysis</u> (2nd. ed. Horper & Row, New York, Evanston & London And John Weatherhill, Inc., Tokyo) p. 792

$$Y_1 = -27.95880 + 4.69186X_5 - 1.36585X_3 + 2.35724X_1$$

$$a'' = -27.95880$$

$$b_1'' = 4.69186$$

$$b_2' = -1.3685$$

$$b_3 = 2.35724$$

F - value for analysis of variance = 227.228

Standard error of estimate = 3.490

ในขั้นที่ 3 นี้ X_1 เข้ามาช่วย explained Y_1 ต่อจากขั้นที่ 2 ซึ่งทำให้ค่า standard error of estimate ลดลงไปอีก ฉะนั้นจะทำการเลือกตัวแปรต่อไปในขั้นที่ 4

ขั้นที่ 4 เครื่องคอมพิวเตอร์เลือก X_2 - จำนวนบ้านในเขตนครหลวง เป็นตัวแปรตัวที่ 4 ที่เข้ามาช่วย explained Y_1 ซึ่งจะได้โมเดลมีลักษณะดังนี้

$$Y_1 = -157.15234 + 4.29610X_5 - 0.29184X_3 + 17.36565X_1 - 10.11832X_2$$

$$a''' = -157.15234$$

$$b_1''' = 4.29610$$

$$b_2'' = -0.29184$$

$$b_3' = 13.36565$$

$$b_4 = -10.11832$$

F - value for analysis of variance = 342.408

Standard error of estimate = 2.467

เมื่อ X_2 เข้ามาช่วย explained Y_1 เพิ่มอีกตัวหนึ่ง จะทำให้ค่า standard error of estimate ลดลงจากขั้นที่ 3 อีก และค่า F เพิ่มมากขึ้นอีก แสดงว่าสมการในขั้นที่ 4 นี้จะให้ค่าประมาณ Y_1 แตกต่างจากค่าจริง Y_1 น้อยลงไปอีก ต่อไปก็จะเหลือตัวแปรอิสระอีกตัวเดียว ซึ่งจะนำมาใส่ในโมเดลต่อไป

ขั้นที่ 5 ตัวแปรอิสระตัวสุดท้าย คือ X_4 - G.D.P ในเขตนครหลวง เมื่อนำมาช่วย explained Y_1 จะได้โมเดลดังนี้

$$Y_1 = -154.19664 + 4.22509X_5 - 0.34010X_3 + 17.49355X_1 - 10.35210X_2 + 0.29383X_4$$

$$a^{IV} = -154.19664$$
$$b_1^{IV} = 4.22509$$
$$b_2^{'''} = -0.34010$$
$$b_3^{''} = 17.49355$$
$$b_4^{'} = -10.35210$$
$$b_5 = .29383$$

F - value for analysis of variance = 205.644

Standard error of estimate = 2.848

จะเห็นว่า ค่า standard error of estimate เพิ่มขึ้นจากขั้นที่ 4 แสดงว่าเมื่อนำ X_4 เข้ามาในโมเดล X_4 ไม่ได้ช่วย explained Y_1 เพราะทำให้ standard error of estimate เพิ่มขึ้นกว่าเดิมอีก ฉะนั้นเราจะไม่นำ X_4 เข้ามาใช้ในโมเดลของเรา แสดงว่าค่า G.D.P. ซึ่งเราใช้เป็นค่าแทนสภาพเศรษฐกิจไม่ significant พอที่จะเป็นตัวแปรอิสระอีก 1 ตัวในโมเดล ควรจะรวมอยู่ใน residual

ดังนั้น เราจะใช้โมเดลในขั้นที่ 4 สำหรับการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทในเขตนครหลวง ซึ่งมีลักษณะโมเดลดังนี้

$$Y_1 = -157.15234 + 4.29610X_5 - 0.29184X_3 + 17.36565X_1 - 10.11832X_2$$

เมื่อนำค่าของตัวแปรอิสระทั้ง 4 ตัวคือ X_1, X_2, X_3, X_5 แทนลงไปในโมเดล จะได้ค่าประมาณ Y_1 ตามตารางที่ 2.1.11 นี้

ตารางที่ 2.1.11

แสดงค่าประมาณ $\hat{Y}_1$ โดยวิธี Multiple Linear Regression

ปี	Y_1	$\hat{Y}_1$
2507	35.064	35.742
2508	39.155	39.910
2509	42.060	38.734
2510	43.861	44.443
2511	48.170	49.317
2512	56.395	56.123
2513	66.384	66.475
2514	105.550	108.013
2515	127.221	125.106

Y_1 เป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท (จากตารางที่ 2.1.1)

$\hat{Y}_1$ เป็นค่าประมาณของ Y_1 ซึ่งประมาณจากโมเดลของ multiple linear regression คือ

$$Y_1 = -157.15234 + 4.29610X_5 - 0.29184X_3 + 17.36565X_1 - 10.11832X_2$$

2.2 วิธีการประมาณจำนวนครั้งที่เรียกโทรศัพท์ในเขตนครหลวง

2.2.1 การประมาณจำนวนครั้งที่เรียกโทรศัพท์ในเขตนครหลวงโดย Trend Method แบ่งการประมาณตามประเภทผู้เช่าดังนี้

1) การประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท หมายถึงการประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมของประเภทร้านค้า บ้านพัก พิเศษ ราชการ สาธารณะ ทศท. ซึ่งข้อมูลแสดงไว้ในตารางที่ 2.2.1

ตารางที่ 2.2.1

แสดงข้อมูลของจำนวนครั้งที่เรียกโทรศัพท์ประเภทต่าง ๆ เป็นรายปี

(ล้านครั้ง)

ปี	U_1	U_2	U_3	U_4	U_5	Y'_2	U_6	Y_2
2507	59.109	14.940	3.710	16.220	5.615	99.594	.675	100.269
2508	68.922	17.649	4.405	18.381	6.455	115.812	.900	116.712
2509	76.493	20.086	4.887	19.139	6.602	127.207	1.422	128.629
2510	85.473	22.104	5.705	19.671	7.122	140.075	1.538	141.613
2511	89.059	23.194	5.949	20.357	6.745	145.304	1.851	147.155
2512	94.155	24.891	6.362	24.012	7.222	156.642	1.966	158.608
2513	91.042	25.871	6.932	25.029	7.983	156.857	2.397	159.254
2514	112.712	41.867	7.107	25.960	8.287	195.933	2.478	198.411
2515	163.583	71.436	9.989	29.423	12.081	286.512	3.015	289.527

U_1 = จำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า

U_2 = จำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก

U_3 = จำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ

U_4 = จำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการ

U_5 = จำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ

Y'_2 = จำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท

U_6 = จำนวนครั้งที่เรียกโทรศัพท์ประเภท หศท.

Y_2 = จำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท

Y_2 เป็นจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท

โดย Trend Method จะหาเส้นแนวโน้มทั้งที่เป็นเส้นตรง และไม่เป็นเส้นตรง แล้วพิจารณาเลือกสมการในการคำนวณว่าจะใช้สมการใด

แสดงสมการต่าง ๆ ที่ใช้คำนวณหาค่าประมาณ
ของจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท

สมการ	Standard error of estimate	เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระ
(1) $\hat{Y}_2 = 87.994846 + 18.006233X$	28.612015	77.25
(2) $\hat{Y}_2 = 115.970038 - 5.972503X + 2.997342X^2$	22.218891	88.24
(3) $\hat{Y}_2 = (99.92)(1.112)^X$	25.008260	82.62
(4) $\hat{Y}_2 = 80.030530 + (27.571123)(1.252)^X$	22.266451	88.18
(5) $\hat{Y}_2 = (46.93)(2.271)^{1.090^X}$	23.836844	86.46
(6) $\hat{Y}_2 = \dfrac{1}{-7728.430414 + (17201.900997)(.9568)^X}$	24.768418	82.95

การพิจารณาตัดสินใจว่า จะใช้สมการใดในการคำนวณหาค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท ก็ยังคงใช้เหตุผลเดียวกันกับการตัดสินใจเลือกสมการในการคำนวณหาค่าประมาณของจำนวนเลขหมายโทรศัพท์รวมทุกประเภท คือ สมการใดให้ค่า standard error of estimate ต่ำสุดและเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระสูงสุดแล้ว จะเลือกใช้สมการนั้นในการประมาณ สมการที่ (2) second - degree polynomial ให้ค่า standard error of estimate เท่ากับ 22.218891 ซึ่งต่ำสุด และ 88.24 % ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ดังนั้น เราจะเลือกใช้สมการ (2) เป็นสมการในการคำนวณหาเส้นแนวโน้มของจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท ซึ่งจะให้ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภทที่ใกล้เคียงกับค่าจริงที่สุด ตามตารางที่ 2.2.2 นี้

ตารางที่ 2.2.2

แสดงการเปรียบเทียบค่าประมาณจากสมการ ($\hat{Y}_2$) กับค่าจริง (Y_2)

(ล้านครั้ง)

ปี	X	Y_2	$\hat{Y}_2$
2507	0	100.269	115.970038
2508	1	116.712	112.994877
2509	2	128.629	116.014400
2510	3	141.613	125.028607
2511	4	147.155	140.037498
2512	5	158.608	161.041073
2513	6	159.254	188.039332
2514	7	198.411	221.032275
2515	8	289.527	260.019902

X กำหนดขึ้นเป็นรายปี โดยให้ 0,1,2,...,8 ตามลำดับ

$\hat{Y}_2$ ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท ซึ่งประมาณจากสมการ second - degree polynomial คือ

$$Y_2 = a + b + cx^2 + e$$

โดยวิธี least squares จะได้ normal equation ตามโมเดล 4 หน้า 10 ที่สามารถหาค่า a,b,c ได้ดังนี้

$$1440.178 = 9a + 36b + 204c$$

$$6841.086 = 36a + 204b + 1296c$$

$$42210.436 = 204a + 1296b + 8772c$$

จาก 3 สมการที่มี unknown parameters 3 ตัว ก็สามารถ solve หาค่า a,b,c ได้คือ

$$a = 115.970038$$

$$b = -5.972503$$

$$c = 2.997342$$

เพราะฉะนั้น สมการ second-degree polynomial ที่ใช้ประมาณค่าจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท จะมีรูปสมการดังนี้

$$\hat{Y}_2 = 115.970038 - 5.972503X + 2.997342X^2$$

ทดสอบสมมติฐานเกี่ยวกับ ค่า b,c ในสมการโดยทำ analysis of variance เหมือนกับตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_0 : b = 0$

$F = 23.762937$ มี(1,7) degree of freedom

ค่า F มากกว่า F (1,7) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject $H_0: b = 0$ นั่นคือ $b \neq 0$

$H_0 : c = 0$

$F = 5.607806$ มี (1,6) degree of freedom

ค่า F มากกว่า F (1,6) จากตาราง F - distribution ที่ $\alpha = .10$ ดังนั้น reject $H_0 : c = 0$ นั่นคือ $c \neq 0$ แสดงว่า ข้อมูลจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท มีแนวโน้มเป็นเส้นโค้งของ second-degree polynomial ดูกราฟแสดงภาพประกอบที่ 8 ในภาคผนวก ก

2) <u>การประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า</u> ประมาณจากข้อมูลของจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า ซึ่งเกิดจากการใช้โทรศัพท์ของเลขหมายโทรศัพท์ประเภทร้านค้าเป็นรายปี ตามที่แสดงไว้ในตารางที่ 2.2.1 โดยนำมาคำนวณหาเส้นแนวโน้มเพื่อหาค่าประมาณจากเส้นแนวโน้มนั้น ทั้งนี้คำนวณหาทั้งที่

เป็นเส้นตรง และไม่เป็นเส้นตรง แล้วพิจารณาตัดสินใจว่าจะเลือกใช้สมการใดในการประมาณ ซึ่งใช้หลักเกณฑ์เดียวกันกับที่กล่าวไว้แล้วในตอนที่ (1) ของข้อ 2.1.1 จากการคำนวณหาสมการต่าง ๆ ทั้ง 6 สมการ คือ

1. สมการเส้นตรง
2. สมการ second - degree polynomial
3. สมการ simple exponential
4. สมการ modified exponential
5. สมการ gompertz curve
6. สมการ logistic curve

ปรากฏว่า สมการ second - degree polynomial ให้ค่า standard error of estimate เท่ากับ 13.315598 ซึ่งต่ำสุด และ 85.72 % ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ดังนั้น เราจะตัดสินใจเลือกสมการ second - degree polynomial หาค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า เพื่อที่จะได้ค่าประมาณที่ใกล้เคียงกับค่าจริงที่สุด ดังตาราง 2.2.3 นี้

ตารางที่ 2.2.3

แสดงค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์แยกประเภทโดย Trend Method

(ล้านครั้ง)

ปี	X	$\hat{U}_1$	$\hat{U}_2$	$\hat{U}_3$	$\hat{U}_4$	$\hat{U}_5$	$\hat{Y}'_2$
2507	0	67.705109	20.416201	3.892	17.216756	6.332966	115.265351
2508	1	67.403626	16.507382	4.316	17.910806	6.120675	112.014270
2509	2	69.983743	15.205821	4.787	18.830217	6.131810	114.758267
2510	3	75.445460	16.511518	5.309	19.974991	6.366371	123.497342
2511	4	83.788777	20.424473	5.887	21.345127	6.824358	138.231495
2512	5	95.013694	26.944686	6.528	22.940624	7.505771	158.960726
2513	6	109.120211	36.072157	7.240	24.761484	8.410610	185.685035
2514	7	126.108328	47.806886	8.028	26.807706	9.538875	218.404422
2515	8	145.978045	62.148873	8.902	29.079289	10.890566	257.118887

X กำหนดขึ้นตามปี โดยให้เท่ากับ 0, 1, 2, ..., 8 ตามลำดับ

$\hat{U}_1$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า

$\hat{U}_2$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก

$\hat{U}_3$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ

$\hat{U}_4$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการ

$\hat{U}_5$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ

$\hat{Y}'_2$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท

$\hat{U}_1$ ในตาราง 2.2.3 เป็นค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภท ร้านค้า ซึ่งประมาณจาก

$$U_1 = a + bx + cx^2 + e$$

โดยวิธี least squares หาค่า a,b,c ได้

$a = 67.705109$

$b = -1.742283$

$c = 1.4408$

ดังนั้นสมการที่ให้ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า ตามตารางที่ 2.2.3 จะมีรูปสมการดังนี้

$$\hat{U}_1 = 67.705109 - 1.742283X + 1.4408X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b,c โดยทำ analysis of variance ตามตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_0 : b = 0$

$F = 23.600844$ มี (1,7) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่า ค่า F (1,7) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject $H_0 : b = 0$ แสดงว่า $b \neq 0$

$H_0 : c = 0$

$F = 29.847893$ มี (1,6) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่าค่า F (1,6) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject $H_0 : c = 0$ นั่นคือ $c \neq 0$ แสดงว่า ข้อมูลของจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า มีแนวโน้มเป็น second - degree polynomial ซึ่งจะดูภาพประกอบด้วยจากกราฟแสดงภาพที่ 9 ในภาคผนวก ก

3) <u>การประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก</u> เป็นการประมาณจากข้อมูลจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก ซึ่งเกิดจากการใช้โทรศัพท์ของเลขหมายโทรศัพท์ประเภทบ้านพัก ดังแสดงไว้ในตารางที่ 2.2.1

U_2 เป็นจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก

วิธีประมาณ เราจะหาเส้นแนวโน้มของ U_2 ทั้งที่เป็นเส้นตรงและไม่เป็นเส้นตรง จาก 6 สมการที่กล่าวไว้ในตอนที่ (2) ข้อ 2.2.1 และการพิจารณาว่าจะใช้สมการใดนั้น ก็คงใช้หลักเกณฑ์เดียวกันกับการประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท ดังนั้น จากการคำนวณทั้ง 6 สมการ ปรากฏว่า สมการ modified exponential

$$\hat{U}_2 = 15.571832 + (1.174121)(1.579)^X$$

ให้ค่า standard error of estimate ต่ำสุดคือ 5.313236 และ 92.03 เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ดังนั้น สมการ modified exponential นี้ จะให้ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก ($\hat{U}_2$) ที่ใกล้เคียงกับค่าจริง แต่ปรากฏว่าเมื่อใช้หาค่าพยากรณ์ในอนาคตโดยวิธี extrapolation แล้วจะให้ค่าพยากรณ์มากเกินความเป็นจริง คือ ค่าพยากรณ์จำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก ที่เพิ่มขึ้นในแต่ละปีจะเพิ่มขึ้นทีละเท่าตัว นอกจากนี้ค่าพยากรณ์จำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพักในแต่ละปี มากกว่าค่าพยากรณ์จำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภทซึ่งเป็นไปไม่ได้ ดังนั้น เราจึงเลือกใช้สมการอื่นที่ให้ค่า standard error of estimate ต่ำสุด และเปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระสูงสุด ในทุกสมการยกเว้นสมการ modified exponential ซึ่งได้แก่ สมการ second - degree polynomial

$$\hat{U}_2 = 20.416201-5.212448X + 1.303629X^2$$

ให้ค่า standard error of estimate เท่ากับ 7.346978 และ 86.94 เปอร์เซนต์ของการกระจายของตัวแปรไม่อิสระ ซึ่งจะให้ค่าประมาณ ($\hat{U}_2$) ใกล้เคียงกับค่าจริง (U_2) เช่นกัน และเมื่อหาค่าพยากรณ์ในอนาคตแล้ว จะได้ตัวเลขค่าพยากรณ์ที่เหมาะสมด้วย

$\hat{U}_2$ จากตาราง 2.2.3 คือ ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพักจากสมการ second degree polynomial

$$U_2 = a + bx + cx^2 + e$$

โดยวิธี least squares หาค่า a,b,c ได้

$$a = 20.416201$$

$$b = -5.212448$$

$$c = 1.303629$$

เพราะฉะนั้น สมการ second - degree polynomial ที่ใช้ประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก มีรูปสมการดังนี้

$$\hat{U}_2 = 20.416201-5.212448X + 1.303629X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b,c โดยทำ analysis of variance ตามตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_o : b = 0$

$F = 13.487286$ มี (1,7) degree of freedom

ค่า F ที่คำนวณได้จากการทำ analysis of variance มากกว่าค่า F (1,7) จากตาราง F distribution ที่ค่า α = .10, .05, .025, .010 ดังนั้น reject $H_o : b = 0$ แสดงว่า $b \neq 0$

$H_o : c = 0$

$F = 9.699272$ มี (1,6) degree of freedom

ค่า F ที่คำนวณได้จากการทำ analysis of variance มากกว่าค่า F (1,6) จากตาราง F - distribution ที่ค่าของ α = .10, .05, .025 ดังนั้น reject $H_o : c = 0$ นั่นคือ $c \neq 0$ แสดงว่าสมการ

$$\hat{U}_2 = 20.416201-5.212448X + 1.303629X^2$$

ก็ยังเหมาะสมที่จะใช้ประมาณค่าจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก ดูกราฟแสดงภาพที่ 10 ในภาคผนวก ก

4) <u>การประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ</u> เป็นการประมาณจากข้อมูลจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ ซึ่งเกิดจากการใช้โทรศัพท์ของเลขหมายโทรศัพท์ประเภทพิเศษ ดังแสดงไว้ในตารางที่ 2.2.1

U_3 เป็นจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ

วิธีประมาณ คือหาเส้นแนวโน้มของ U_3 ทั้งที่เป็นเส้นตรงและไม่เป็นเส้นตรง รวม 6 สมการ การพิจารณาตัดสินใจว่าจะใช้สมการใดนั้น คงใช้หลักเกณฑ์เดียวกันกับที่ได้กล่าวแล้วในตอนที่ (2) ข้อ 2.2.1 ดังนั้น สมการที่เหมาะสมที่จะใช้ในการประมาณ U_2 ได้แก่ Simple exponential

$$U_3 = (3.892)(1.109)^X$$

ซึ่งให้ค่า standard error of estimate ต่ำสุด คือ .581124 และ 91.29 % ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ดังนั้น สมการ simple exponential นี้จะให้ค่าประมาณ $\hat{U}_3$ ใกล้เคียงกับค่าจริง U_3 ที่สุด

$\hat{U}_3$ จากตาราง 2.2.3 เป็นค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ ซึ่งประมาณจากเส้นแนวโน้มที่มีลักษณะเป็น simple exponential

$$U_3 = ab^X$$

$$\log U_3 = \log a + X\log b$$

โดยวิธี least squares ได้ normal equation ซึ่งแก้สมการ ได้ค่า

$$\log a = .590239$$

$$\log b = .044907$$

ดังนั้น simple exponential ที่อยู่ในรูป linear logarithm จะมีรูปสมการดังนี้

$$\log \hat{U}_3 = .590239 + .044907X$$

เปิดตาราง anti - log ได้ค่า a,b ดังนี้

$$\hat{U}_3 = (3.892)(1.109)^X$$

ทดสอบสมมติฐานเกี่ยวกับค่า log b โดยทำ analysis of variance ตามตารางที่ 2.1.6

H_o : log b = 0

$F = 114.907882$ มี (1,7) degree of freedom

ค่า F ที่คำนวณได้มากกว่า F (1,7) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject H_0 นั่นคือ $\log b \neq 0$ แสดงว่า สมการ simple exponential เหมาะสมที่จะใช้ประมาณค่าจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ เพราะให้ค่าประมาณ $\hat{U}_3$ ที่ใกล้เคียงกับค่าจริง U_3 ที่สุด หรืออีกนัยหนึ่งคือ ข้อมูลของจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษมีแนวโน้มในลักษณะ linear logarithm ดังกราฟแสดงภาพที่ 11 ในภาคผนวก ก

5) <u>การประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการ</u> ประมาณจากข้อมูลของจำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการเป็นรายปี ซึ่งเกิดจากการใช้โทรศัพท์ของเลขหมายโทรศัพท์ประเภทราชการ ตามที่แสดงไว้ในตารางที่ 2.2.1

U_4 เป็นจำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการ

วิธีประมาณ คือคำนวณหาเส้นแนวโน้มของ U_4 ในรูปสมการเส้นตรงและไม่ใช่เส้นตรง รวม 6 สมการ ตามที่กล่าวไว้แล้วในตอนที่ (2) ข้อ 2.2.1 การพิจารณาเลือกใช้สมการใดนั้น ก็ใช้เหตุผลเกี่ยวกับการประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทอื่น ๆ ที่ได้กล่าวมาแล้ว ดังนั้น สมการที่เหมาะสมในการประมาณคือ second - degree **polynomial**

$$\hat{U}_4 = 17.216756 + .5813687X + .112681X^2$$

ซึ่งให้ค่า standard error of estimate ต่ำสุด คือ .768781 และ 97.45 % ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ฉะนั้น **สมการ** second - degree **polynomial** จะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงที่สุด

$\hat{U}_4$ จากตาราง 2.2.3 คือ ค่าประมาณจำนวนครั้งโทรศัพท์ประเภทราชการ ซึ่งคำนวณจาก second - degree polynomial

$$U_4 = a + bx + cx^2 + e$$

โดยวิธี least squares แก้สมการหาค่า a,b,**c** ได้

$a = 17.216756$

$b = .5813687$

$c = .112681$

ดังนั้น สมการ second - degree polynomial ที่ใช้คำนวณหาเส้นแนวโน้ม ของ U_4 ที่จะให้ค่าประมาณ $\hat{U}_4$ จะอยู่ในรูป

$$\hat{U}_4 = 17.216756 + .5813687X + .112681X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b,c โดย analysis of variance ดังตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

H_o : b = 0

$F = 123.733201$ มี (1,7) degree of freedom

ค่า F ที่คำนวณได้มากกว่าค่า F (1,7) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject H_o : b = 0 นั่นคือ b ≠ 0

H_o : c = 0

$F = 6.50238$ มี (1,6) degree of freedom

ค่า F ที่คำนวณได้มากกว่า F (1,6) จากตาราง F - distribution ที่ α = .10, .05 เพราะฉะนั้น reject Ho นั่นคือ c ≠ 0 แสดงว่า ข้อมูลจำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการมีแนวโน้มในรูป second - degree polynomial จึงเหมาะสมที่จะใช้สมการดังกล่าวคำนวณหาค่าประมาณ ดูกราฟแสดงภาพประกอบที่ 12 ในภาคผนวก ก

6) <u>การประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ</u> ประมาณจากเส้นแนวโน้มของจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ โดยคำนวณจากข้อมูลจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะที่เกิดจากการใช้โทรศัพท์ของเลขหมายโทรศัพท์ประเภทสาธารณะ ดังแสดงไว้ในตารางที่ 2.2.1

U_5 เป็นจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ

การหาเส้นแนวโน้มของ U_5 จะคำนวณหาทั้งสมการที่เป็นเส้นตรงและไม่ใช่เส้นตรง รวม 6 สมการนั้น แล้วพิจารณาเลือกว่าจะสมการใดที่เหมาะสมกับข้อมูล U_5 โดยหลักเกณฑ์เดียวกับที่กล่าวในตอน (2) ข้อ 2.2.1 ซึ่งได้แก่ สมการ second - degree polymonial

$$\hat{U}_5 = 6.332966 - .324004X + .111713X^2$$

ให้ค่า standard error of estimate ต่ำสุด คือ .882545 และ 83.33 % ของการกระจายของตัวแปรไม่อิสระซึ่งสูงสุด ดังนั้น สมการ second-degree polynomial นี้จะให้ค่าประมาณที่ใกล้เคียงกับค่าจริง U_5 ที่สุด

$\hat{U}_5$ จากตารางที่ 2.2.3 เป็นค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ที่ประมาณจากสมการ second-degree polynomial

$$U_5 = a + bX + cX^2 + e$$

โดยวิธี least squares แก้สมการหาค่า a,b,c ได้ดังนี้

$a = 6.332966$

$b = -.324004$

$c = .111713$

เพราะฉะนั้น สมการ second-degree polynomial ที่ใช้ประมาณค่าจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ จะมีรูปสมการดังนี้

$$\hat{U}_5 = 6.332966 - .324004\ X + .111713\ X^2$$

ทดสอบสมมติฐานเกี่ยวกับ ค่า b,c โดย analysis of variance ตามตารางที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_o : b = 0$

$F = 10.176266$ มี (1,7) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่า F (1,7) จากตาราง F-distribution ที่ค่า $\alpha = .10,\ .05$ และ $.025$ ดังนั้น reject H_o นั่นคือ $b \neq 0$

$H_o : c = 0$

$F = 8.670670$ มี (1,6) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่า F (1,6) จากตาราง F - distribution ที่ค่า $\alpha = .10,\ .05$ ดังนั้น reject H_o นั่นคือ $c \neq 0$ แสดงว่า สมการ second-degree polynomial ใช้สำหรับประมาณค่าจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะได้ เพราะข้อมูลมีเส้นแนวโน้มในลักษณะ second-degree polynomial ดูภาพแสดงกราฟประกอบที่ 13 ในภาคผนวก ก

7) <u>การประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท.</u> ประมาณจาก เส้นแนวโน้มของจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท. ซึ่งเกิดจากการใช้โทรศัพท์ของเลขหมายโทรศัพท์ประเภททศท. เนื่องจากการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท. ในตอนที่ (7) ข้อ 2.1.1 เรามิได้ประมาณจากข้อมูลของจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าประเภททศท. โดยตรง แต่ประมาณจำนวนรวมของเลขหมายโทรศัพท์ที่มีผู้เช่า 5 ประเภท แล้วนำไปหักจากจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท ดังนั้น การประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท. เราจะใช้วิธีประมาณเช่นเดียวกัน คือประมาณจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท แล้วนำไปหักจากจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท ก็จะได้ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท. ข้อมูลจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท แสดงไว้ในตารางที่ 2.2.1 คือใน column Y_2'

การคำนวณหาเส้นแนวโน้มของ Y_2' จะคำนวณทั้งสมการเส้นตรงและไม่ใช่เส้นตรงรวม 6 สมการ แล้วพิจารณาเลือกสมการที่ให้ค่าประมาณใกล้เคียงกับค่าจริงที่สุด ซึ่งมีหลักเกณฑ์ตามที่กล่าวแล้วในตอนที่ (2) ข้อ 2.2.1 สมการดังกล่าว คือ second - degree polynomial

$$\hat{Y}_2' = 115.265351 - 6.24862\,X + 2.997539\,X^2$$

ให้ค่า standard error of estimate ต่ำสุด คือ 19.785492 และ 96.96 % ของการกระจายของตัวแปรอิสระซึ่งสูงสุด ดังนั้น จะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงมากที่สุด

$\hat{Y}_2'$ จากตาราง 2.2.3 คือ ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท ซึ่งประมาณจากสมการ second - degree polynomial

$$Y_2' = a + bX + cX^2 + e$$

แก้สมการโดย least square method หาค่า a,b,c ได้ดังนี้

$a = 115.265351$

$b = -6.24862$

$c = 2.997539$

ดังนั้น สมการ second - degree polynomial ที่ใช้ประมาณค่า $\hat{Y}'_2$ ในตารางที่ 2.2.3 มีรูปสมการดังนี้

$$\hat{Y}'_2 = 115.265351-6.24862X + 2.997539X^2$$

ทดสอบสมมติฐานเกี่ยวกับค่า b,c โดย analysis of variance ตามตาราง ที่ 2.1.3 และตารางที่ 2.1.4 ตามลำดับ

$H_o : b = 0$

$F = 23.110602$ มี (1,7) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่า F (1,7) จากตาราง F - distribution ทุกค่าของ α ดังนั้น reject H_o แสดงว่า $b \neq 0$

$H_o : c = 0$

$F = 143.232739$ มี (1,6) degree of freedom

ค่า F ที่คำนวณได้นี้มากกว่า F (1,6) จากตาราง F - distribution ทุกค่าของ α เพราะฉะนั้น reject H_o นั่นคือ $c \neq 0$ แสดงว่าข้อมูลจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท มีแนวโน้มเป็น second - degree polynomial ดังนั้นเมื่อใช้สมการนี้ประมาณค่า จะให้ค่าประมาณที่ใกล้เคียงกับค่าจริงที่สุด ดูภาพแสดงกราฟประกอบที่ 14 ในภาคผนวก ก

เมื่อได้ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภทแล้ว ก็นำไปหักจากค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภทจากตารางที่ 2.2.2 column $\hat{Y}'_2$ จะได้ค่าประมาณจำนวนครั้งที่เรียก โทรศัพท์ประเภททศท. ดังตารางที่ 2.2.4 ต่อไปนี้

ตารางที่ 2.2.4

แสดงค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท.

(ล้านครั้ง)

ปี	$\hat{Y}_2$	$\hat{Y}'_2$	$\hat{Y}_2 - \hat{Y}'_2$
2507	115.970038	115.265351	.704687
2508	112.994877	112.014270	.980607
2509	116.014400	114.758267	1.256133
2510	125.028607	123.497342	1.531265
2511	140.037498	138.231495	1.806003
2512	161.041073	158.960726	2.080347
2513	188.039332	185.685035	2.354297
2514	221.032275	218.404422	2.627853
2515	206.019902	257.118887	2.901015

$\hat{Y}'_2$ ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท (จากตารางที่ 2.2.2)

$\hat{Y}'_2$ ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท (จากตารางที่ 2.2.3)

$\hat{Y}_2 - \hat{Y}'_2$ ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท.

วิธีการปรับตัวเลขโดย correction factors

ด้วยเหตุผลเดียวกับที่กล่าวในตอนที่ (8) ข้อ 2.1.1 จึงต้องปรับตัวเลขค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์แต่ละประเภท ยกเว้นประเภททศท. ให้ได้ผลรวมเท่ากับค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท โดยการทำ correction factors ระหว่าง ผลรวมของค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ 5 ประเภท กับค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท ตามตารางที่ 2.2.5 ดังนี้

ตารางที่ 2.2.5

แสดง correction factors ของจำนวนครั้งที่เรียกโทรศัพท์

(ล้านครั้ง)

ปี	$\hat{U}_1$	$\hat{U}_2$	$\hat{U}_3$	$\hat{U}_4$	$\hat{U}_5$	$\hat{Y}''_2$	$\hat{Y}'_2$	C
2507	67.705109	20.416201	3.892	17.216756	6.332966	115.563032	115.265351	.99742408
2508	67.403626	16.507382	4.316	17.910806	6.120675	112.258489	112.014270	.99782449
2509	69.983743	15.205821	4.787	18.830217	6.131810	114.938591	114.758267	.99843113
2510	75.445460	16.511518	5.309	19.974991	6.366371	123.607340	123.497342	.99911010
2511	83.788777	20.424473	5.887	21.345127	6.824358	138.269735	138.231495	.99972344
2512	95.013694	26.944686	6.528	22.940624	7.505771	158.932775	158.960726	1.00017587
2513	109.120211	36.072157	7.240	24.761484	8.410610	185.604462	185.685035	1.00043411
2514	126.108328	47.806886	8.028	26.807706	9.538875	218.289795	218.404422	1.00052511
2515	145.978045	62.148873	8.902	29.079289	10.890566	256.998773	257.118887	1.00046737

$\hat{U}_1$ = ค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า (จากตารางที่ 2.2.3)

$\hat{U}_2$ = ค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก (จากตารางที่ 2.2.3)

$\hat{U}_3$ = ค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ (จากตารางที่ 2.2.3)

$\hat{U}_4$ = ค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการ (จากตารางที่ 2.2.3)

$\hat{U}_5$ = ค่าประมาณของจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ(จากตารางที่ 2.2.3)

$\hat{Y}_2'' = \hat{U}_1 + \hat{U}_2 + \hat{U}_3 + \hat{U}_4 + \hat{U}_5$

$\hat{Y}_2'$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวม 5 ประเภท (จากตารางที่ 2.2.3)

C = correction factors ได้จาก Y_2'/ Y_2''

ค่า correction factors ที่ได้จากตารางที่ 2.2.5 นี้นำไปปรับค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า, บ้านพัก, พิเศษ, ราชการ, สาธารณะ โดยทำค่า C แต่ละปี คูณกับค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์แต่ละประเภทในตารางที่ 2.2.5 เป็นรายปี จะให้ค่าประมาณใหม่ในตารางที่ 2.2.6 ซึ่งผลรวมของค่าประมาณใหม่ของจำนวนครั้งที่เรียกโทรศัพท์ 5 ประเภท กับค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท. จะเท่ากับ ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท

ตารางที่ 2.2.6

แสดงค่าประมาณใหม่ของจำนวนครั้งที่เรียกโทรศัพท์ 5 ประเภท โดยวิธี correction factor

(ล้านครั้ง)

ปี	$\hat{U}'_1$	$\hat{U}'_2$	$\hat{U}'_3$	$\hat{U}'_4$	$\hat{U}'_5$	$\hat{U}_6$	$\hat{Y}_2$
2507	67.530706	20.363610	3.881975	17.172407	6.316653	.704687	115.970038
2508	67.256989	16.471470	4.306611	17.871841	6.107359	.980607	112.994877
2509	69.873948	15.181965	4.779490	18.800674	6.122190	1.256133	116.014400
2510	75.378321	16.496824	5.304276	19.957215	6.360706	1.531265	125.028607
2511	83.765604	20.418824	5.885372	21.339224	6.822471	1.806003	140.037498
2512	95.030404	26.949425	6.529148	22.944658	7.507091	2.080347	161.041073
2513	109.167581	36.087816	7.243143	24.772234	8.414261	2.354297	188.039332
2514	126.174549	47.831990	8.032216	26.821783	9.543884	2.627853	221.032275
2515	146.046271	62.177919	8.906161	29.092880	10.895656	2.901015	260.019902

$\hat{U}_1'$ = ค่าประมาณใหม่ของจำนวนครั้งที่เรียกโทรศัพท์ประเภทร้านค้า โดยวิธี correction factor

$\hat{U}_2'$ = ค่าประมาณใหม่ของจำนวนครั้งที่เรียกโทรศัพท์ประเภทบ้านพัก โดยวิธี correction factor

$\hat{U}_3'$ = ค่าประมาณใหม่ของจำนวนครั้งที่เรียกโทรศัพท์ประเภทพิเศษ โดยวิธี correction factor

$\hat{U}_4'$ = ค่าประมาณใหม่ของจำนวนครั้งที่เรียกโทรศัพท์ประเภทราชการ โดยวิธี correction factor

$\hat{U}_5'$ = ค่าประมาณใหม่ของจำนวนครั้งที่เรียกโทรศัพท์ประเภทสาธารณะ โดยวิธี correction factor

$\hat{U}_6'$ = ค่าประมาณจำนวนครั้งที่เรียกโทรศัพท์ประเภททศท. จากตารางที่ 2.2.4 column $\hat{Y}_2 - \hat{Y}_2'$

$\hat{Y}_2'$ = $\hat{U}_1' + \hat{U}_2 + \hat{U}_3' + \hat{U}_4' + \hat{U}_5' + \hat{U}_6$ ซึ่งเท่ากับ $\hat{Y}_2$ ในตารางที่ 2.2.2

2.2.2 <u>การประมาณจำนวนครั้งที่เรียกโทรศัพท์โดย Multiple Linear Regression Method</u>

เป็นการประมาณโดยพิจารณาปัจจัยต่าง ๆ ที่เกี่ยวข้องด้วย การสร้างโมเดลคงใช้หลักการของ Step-wise multiple regression เช่นเดียวกับในการประมาณจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท ตัวแปรต่าง ๆ ที่ใช้ในการสร้างโมเดลมีดังนี้

1. ตัวแปรไม่อิสระ $-Y_2$ จำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภทในเขตนครหลวง
2. ตัวแปรอิสระ ได้แก่

Y_1 จำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทในเขตนครหลวง

X_1 จำนวนเครื่องโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทในเขตนครหลวง

X_2 จำนวนประชากรในเขตนครหลวง

X_3 G. D. P ในเขตนครหลวง

X_4 ความหนาแน่นโทรศัพท์ คือจำนวนเครื่องโทรศัพท์ต่อประชากร 1,000 คน

ค่าของตัวแปรไม่อิสระ $-Y_2$ แสดงไว้ในตารางที่ 2.2.1

ค่าของตัวแปรอิสระ Y_1 แสดงไว้ในตารางที่ 2.1.1 สำหรับค่า X_1,X_2,X_3,X_4 แสดงไว้ในตารางที่ 2.2.7 ดังนี้

ตารางที่ 2.2.7

แสดงค่าตัวแปรอิสระ X_1,X_2,X_3,X_4 เป็นรายปี

ปี	X_1 (พันเครื่อง)	X_2 (หมื่นคน)	X_3 (พันล้านบาท)	X_4 (เครื่องต่อ 1,000 คน)
2507	53.686	27.41118	20.9525	19.5854
2508	57.474	28.96352	22.8524	19.8435
2509	63.820	30.01706	25.5871	21.2162
2510	74.818	31.23602	28.9884	23.9524
2511	87.422	32.47339	32.2883	26.9211
2512	103.988	36.46878	37.1255	28.5142
2513	118.809	37.89205	40.8381	31.3545
2514	161.192	39.45254	44.9219	40.8571
2515	190.136	40.96580	49.4141	46.4133

วิธีการสร้างโมเดลโดยใช้หลักการของ Step-wise multiple regression จะให้เครื่องคอมพิวเตอร์ IBM SYSTEM 30 MODEL 40 คำนวณเช่นกัน ซึ่งมีขั้นดำเนินงานเหมือนในข้อ 2.1.2 คือ

ขั้นที่ 1 เครื่องคอมพิวเตอร์คำนวณหา correlation matrix ซึ่งแสดงค่าสัมประสิทธิ์แห่งสหสัมพันธ์ (correlation coefficient) ระหว่างจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท กับตัวแปรอิสระแต่ละตัวดังนี้

	Y_2	Y_1	X_1	X_2	X_3	X_4
Y_2	1,00000	0.95504	0.94956	0.86524	0.89834	0.95235
Y_1			0.98413	0.89426	0.91837	0.98570
X_1				0.95500	0.97274	0.99820
X_2					0.99402	0.94163
X_3						0.96517
X_4						1.0000

จาก correlation matrix ซึ่งเป็น symmetric matrix ค่าสัมประสิทธิ์สหสัมพันธ์ (r) สูงที่สุด คือ Y_1 - จำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภท ดังนั้น $V_1 = r^2_{Y_2Y_1}$ สูงสุดด้วย เครื่องคอมพิวเตอร์จะเลือก Y_1 เข้าในโมเดล เป็น simple linear regression

$$Y_2 = 56.18518 + 1.65735\,Y_1$$

$$a = 56.18518$$

$$b_1 = 1.65735$$

F - value for analysis of variance = 72.633

Standard error of estimate = 17.783

ขั้นที่ 2 เครื่องคอมพิวเตอร์จะเลือกตัวแปรอิสระตัวต่อไปที่ให้ค่า V_k สูงสุด เพื่อนำมาใส่ในโมเดล ซึ่งได้แก่ X_4 - ความหนาแน่นโทรศัพท์ ดังนั้นโมเดลจะเป็น multiple linear regression

$$Y_2 = 31.53888 + 0.99629Y_1 + 2.29864X_4$$

$$a' = 31.53888$$

$$b' = 0.99629$$

$$b_2 = 2.29864$$

F - vaule for analysis of variance = 32.859

Standard error of estimate = 18.739

เมื่อเปรียบเทียบค่า standard error of estimate ในขั้นที่ 1 และขั้นที่ 2 แล้ว ปรากฏว่า ค่า standard error of estimate ในขั้นที่ 2 มากกว่าขั้นที่ 1 แสดงว่าเมื่อเพิ่ม X_4 เข้าไปในสมการ ไม่ได้ช่วยลดค่า residual นั่นคือ X_4 ไม่ได้ช่วย explained Y_2

ขั้นที่ 3 แม้ว่า X_4 จะไม่ได้ช่วย explained Y_2 ก็ตาม เครื่องคอมพิวเตอร์จะไม่ค่านึงถึงแต่จะเลือกตัวแปรอิสระตัวที่ 3 เข้ามาใหม่ เพราะการสร้างโมเดล multiple linear regression โดย step-wise multiple regression ด้วยเครื่องคอมพิวเตอร์นั้นจะเป็นเพียงแต่การเลือกตัวแปรอิสระเข้ามาในสมการตามลำดับความสำคัญมากน้อยที่ตัวแปรอิสระมีต่อตัวแปรไม่อิสระ ดังนั้น ในขั้นที่ 3 นี้ เครื่องคอมพิวเตอร์จะเลือกค่า X_2 - จำนวนประชากรในเขตนครหลวง ซึ่งมีโมเดลดังนี้

$$Y_2 = 56.98180 + 0.65329Y_1 + 4.38477X_4 - 1.88656X_2$$

F - value for analysis of variance = 18.742

Standard error of estimate = 20.281

จะเห็นว่าค่า F - value ลดลงจากขั้นที่ 2 และค่า standard error of estimate สูงขึ้นกว่าขั้นที่ 2 ด้วย แสดงว่า X_2 ไม่ได้ช่วย explained Y_2 เช่นกัน

ขั้นที่ 4 เครื่องคอมพิวเตอร์เลือกได้ X_3 - G.D.P ในเขตนครหลวง โมเดลจะเป็นดังนี้

$$Y_2 = 666.37988 + 3.70086Y_1 - 13.38815X_4 - 32.44385X_2 + 23.85640X_3$$

F - value for analysis of variance = 16.406

Standard error of estimate = 19.019

จะเห็นว่าค่า standard error of estimate ลดลงเพียงเล็กน้อยจากขั้นที่ 3 แต่ก็ยังมากกว่าค่า standard error ในขั้นที่ 1 ดังนั้น X_3 จึงไม่ได้ช่วย explained Y_2 อีกเช่นกัน

ขั้นที่ 5 ตัวแปรอิสระตัวสุดท้าย คือ X_1 - จำนวนเครื่องโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทในเขตนครหลวง เมื่อนำเข้ามาช่วย explained Y_2 จะได้โมเดลดังนี้

$$Y_2 = 1064.63257 + 1.90630Y_1 - 30.62744X_4 - 44.56032X_2 + 25.14574X_3 + 5.12754X_1$$

F - value for analysis of varience = 11.329

Standard error of estimate = 20.548

ปรากฏว่า เมื่อเพิ่ม X_1 เข้าไป ค่า standard error of estimate สูงขึ้นอีก แสดงว่าค่า X_1 มิได้ช่วย explained Y_2 เช่นกัน

ดังนั้น จากการพิจารณาโมเดลในขั้นใดที่ให้ค่า standard error of estimate ต่ำสุด และค่า F สูงสุด ก็จะเลือกโมเดลนั้นใช้หาค่าประมาณ ในที่นี้ได้แก่ โมเดลในขั้นที่ 1 คือ simple linear regression

$$Y_2 = 56.18518 + 1.165735Y_1$$

ที่เป็นเช่นนี้เพราะ ค่า Y_1 ซึ่งเป็นจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่ารวมทุกประเภทนั้น ได้มีการพิจารณาปัจจัยต่าง ๆ ที่นำมาสร้างโมเดลในการประมาณจำนวนครั้งที่เรียกโทรศัพท์ไว้แล้ว ซึ่งได้แก่จำนวนประชากรในเขตนครหลวง G.D.P ในเขตนครหลวง, ความหนาแน่นโทรศัพท์ เป็นต้น ดังนั้นการประมาณจำนวนครั้งที่เรียกโทรศัพท์จึงขึ้นกับจำนวนเลขหมายโทรศัพท์ที่มีผู้เช่าอย่างเดียว เมื่อนำค่า $\hat{Y}_1$ ซึ่งเป็นค่าประมาณของ Y_1 แทนค่าใน model ของ simple linear regression แล้ว จะได้ค่าประมาณของ Y_2 ตามตารางที่ 2.2.8 ดังนี้

ตารางที่ 2.2.8

แสดงการเปรียบเทียบค่า Y_2 กับค่าประมาณ $\hat{Y}_2$

โดยโมเดล simple linear regression

(ล้านครั้ง)

ปี	Y_2	$\hat{Y}_2$
2507	100.269	115.422
2508	116.712	122.330
2509	128.629	120.381
2510	141.613	129.843
2511	147.155	137.921
2512	158.608	149.201
2513	159.254	166.358
2514	198.411	235.201
2515	289.527	263.530

Y_2 เป็นจำนวนครั้งที่เรียกโทรศัพท์รวมทุกประเภท จากตารางที่ 2.2.2

$\hat{Y}_2$ เป็นค่าประมาณของ Y_2 จาก simple linear regression

$Y_2 = 56.18518 + 1.65735Y_1$