บทที่ 4

ผลของการทดสอบโปลิโนเมียลโมเดล

รูปทั่วไปของโปลิโนเมียลโมเดลอันดับที่ m ที่ใช้สำหรับการประมาณค่าของข้อมูลเดิมเขียนเป็นสัญลักษณ์ได้ดังนี้

$$Y = \hat{\alpha}_0 + \hat{\alpha}_1 \phi_1(t) + \hat{\alpha}_2 \phi_2(t) + \ldots\ldots + \hat{\alpha}_m \phi_m(t)$$

โดยที่ $\phi_j(t)$ เป็นโปลิโนเมียลอันดับที่ j เมื่อ j = 0, 1, 2 .... m. ดังนั้น ถ้ากระจาย $\phi_j(t)$ อยู่ในเทอมของ t สมมุติว่าสมการสำหรับการประมาณ คือ

$$Y = \hat{a}_0 + \hat{a}_1 t^2 + \hat{a}_2 t^2 + \ldots\ldots\ldots\ldots + \hat{a}_m t^m$$

โดยที่ $\hat{a}_j$ เป็นตัวคงที่ เมื่อ j = 0, 1, 2 .... m.

วัตถุประสงค์โดยทั่วไปของการวิเคราะห์คือ มุ่งที่ประมาณความสัมพันธ์ระหว่าง y กับ t ด้วยโปลิโนเมียลที่มีอันดับต่ำที่สุดเท่าใดจึงจะเหมาะสม (Fit) กับข้อมูลเดิม หลักเกณฑ์คือสร้างโปลิโนเมียลอันดับที่ m แล้วใช้ F – test พารามิเตอร์ โดยตั้งสมมุติฐานว่า พารามิเตอร์ของตัวผันแปรที่มีอันดับสูงสุดของโปลิโนเมียลเท่ากับศูนย์ และลดหลั่นลงมาจนกระทั่งยอมรับว่าพารามิเตอร์ที่ทดสอบไร้นัยสำคัญ นั่นก็คือ โปลิโนเมียลโมเดลจะมีอันดับต่ำเท่าที่ระดับพารามิเตอร์ไร้นัยสำคัญ

แต่ในการสร้างโปลิโนเมียลโมเดลในที่นี้สามารถทำได้เพียงโปลิโนเมียลอันดับที่ห้า เนื่องจาก λ_j ที่ได้มาจากตารางของ Fisher and Yates in Statistical Tables for Biological, Agricultural and Medical Reserch กำหนดค่า λ_j สูงสุดเท่ากับ λ_5 ฉะนั้น การทดสอบโปลิโนเมียลเพื่อที่จะได้โปลิโนเมียลอันดับต่ำที่สุดที่จะเหมาะสม (Fit) กับข้อมูลเดิม จึงไม่ได้ทดสอบในที่นี้ แต่จะใช้การทดสอบภาวะสารูปสนิทดี (Goodness of Fit) ของข้อมูลเดิมกับโปลิโนเมียลโมเดลอันดับที่สาม, สี่ และห้า ที่สร้างขึ้นมาโดยการทดสอบไคสแควร์

<u>การทดสอบไคสแควร์</u> (χ^2 – test) ไคสแควร์เป็นสถิติที่มีประโยชน์ สามารถใช้ในการทดสอบเกี่ยวกับข้อมูลได้หลายอย่าง เช่น แสดงให้เห็นความแตกต่างของการทดสอบความถี่–

ทวินาม การทดสอบภาวะสารูปสนิทดี (Goodness of Fit) ระหว่างโมเดลทางคณิตศาสตร์ กับข้อมูลที่ได้จากตัวอย่าง การทดสอบความเป็นอิสระของตัวผันแปร เป็นต้น โดยทั่วไปการทดสอบไคสแควร์ประการแรก ต้องตั้งข้อสมมุติฐานภายใต้การสมมุติว่า ความถี่ที่ได้จากการคาดคะเน หรือทางทฤษฎีสามารถคำนวณออกมาได้ ประการที่สอง ความถี่ที่ได้จากตัวอย่างสามารถที่จะสังเกตหรือทดลองได้ นำความถี่ที่ได้มาจากสองประการเปรียบเทียบความแตกต่างกัน พื้นฐานของความแตกต่างเหล่านี้สามารถสร้างเกณฑ์ตัดสินได้ว่า ความถี่ที่ได้จากการคาดคะเน หรือทฤษฎีจะแตกต่างหรือไม่กับความถี่ที่ได้จากการทดลองหรือสังเกต การตัดสินใจว่าจะแตกต่างกันหรือไม่ ขึ้นอยู่กับค่าไคสแควร์ของตัวอย่าง ซึ่งคำนวณได้จากสูตร

$$\chi^2 = \sum_{i=1}^{n} \left[\frac{f_o - f_e}{f_e}\right]^2$$

เมื่อ χ^2 คือ สัญญลักษณ์แทนค่าไคสแควร์

f_o คือ ความถี่ที่ได้จากการสังเกตหรือทดลอง

f_e คือ ความถี่ที่ได้จากการคาดคะเนหรือทฤษฎี

n คือ ขนาดของตัวอย่าง

แต่ในที่นี้จะกล่าวถึงการใช้ไคสแควร์ทดสอบภาวะสารูปสนิทดี (Goodness of Fit) ระหว่างความถี่ที่ได้มาทางทฤษฎีกับความถี่ที่ได้จากข้อมูลเดิม กรรมวิธีในการทดสอบไคสแควร์ทำได้ดังนี้

<u>ตอนที่ 1</u> ตั้งสมมุติฐานว่า

H_o : f_o กับ f_e ไม่แตกต่างกัน คือมีภาวะสารูปสนิทดี หรือ $f_o = f_e$

H_1 : f_o กับ f_e แตกต่างกัน คือไม่มีภาวะสารูปสนิทดี หรือ $f_o \neq f_e$

<u>ตอนที่ 2</u> หาค่าค่าแห่งความเป็นอิสระ (Degree of Freedom) เท่ากับ $(c - 1)(r - 1)$ เมื่อ c คือจำนวนสดมภ์ และ r คือจำนวนแถว เพื่อที่หาค่า χ^2 จากตารางที่กำหนดให้

<u>ตอนที่ 3</u> กำหนดระดับนัยสำคัญตามต้องการ เช่น ระดับนัยสำคัญ 0.01 และ 0.05 เป็นต้น ซึ่งสามารถหาค่า χ^2 จากตาราง χ^2 ณ ระดับนัยสำคัญที่ต้องการ

<u>ตอนที่ 4</u> ทดสอบสถิติ $\chi^2 = \sum_{i=1}^{n} \left[\frac{f_o - f_e}{f_e}\right]^2$

<u>ตอนที่ 5</u> หาเกณฑ์ที่จะตัดสินใจ ได้แก่ ระดับนัยสำคัญ, ค่าแห่งความเป็นอิสระ และ χ^2 จากตาราง

<u>ตอนที่ 6</u> ถ้า χ^2_c ที่ได้จากคำนวณมากกว่า χ^2 ที่ได้จากตาราง จะแสดงว่า H_o มีนัยสำคัญ กล่าวคือ ไม่ยอมรับว่า f_o กับ f_e มีภาวะสารูปสนิทดี ถ้า χ^2 ที่ได้จากการคำนวณน้อยกว่า χ^2 ที่ได้จากตาราง แสดงว่า H_o ไร้นัยสำคัญ กล่าวคือ ยอมรับว่า f_o กับ f_e มีภาวะสารูปสนิทดี

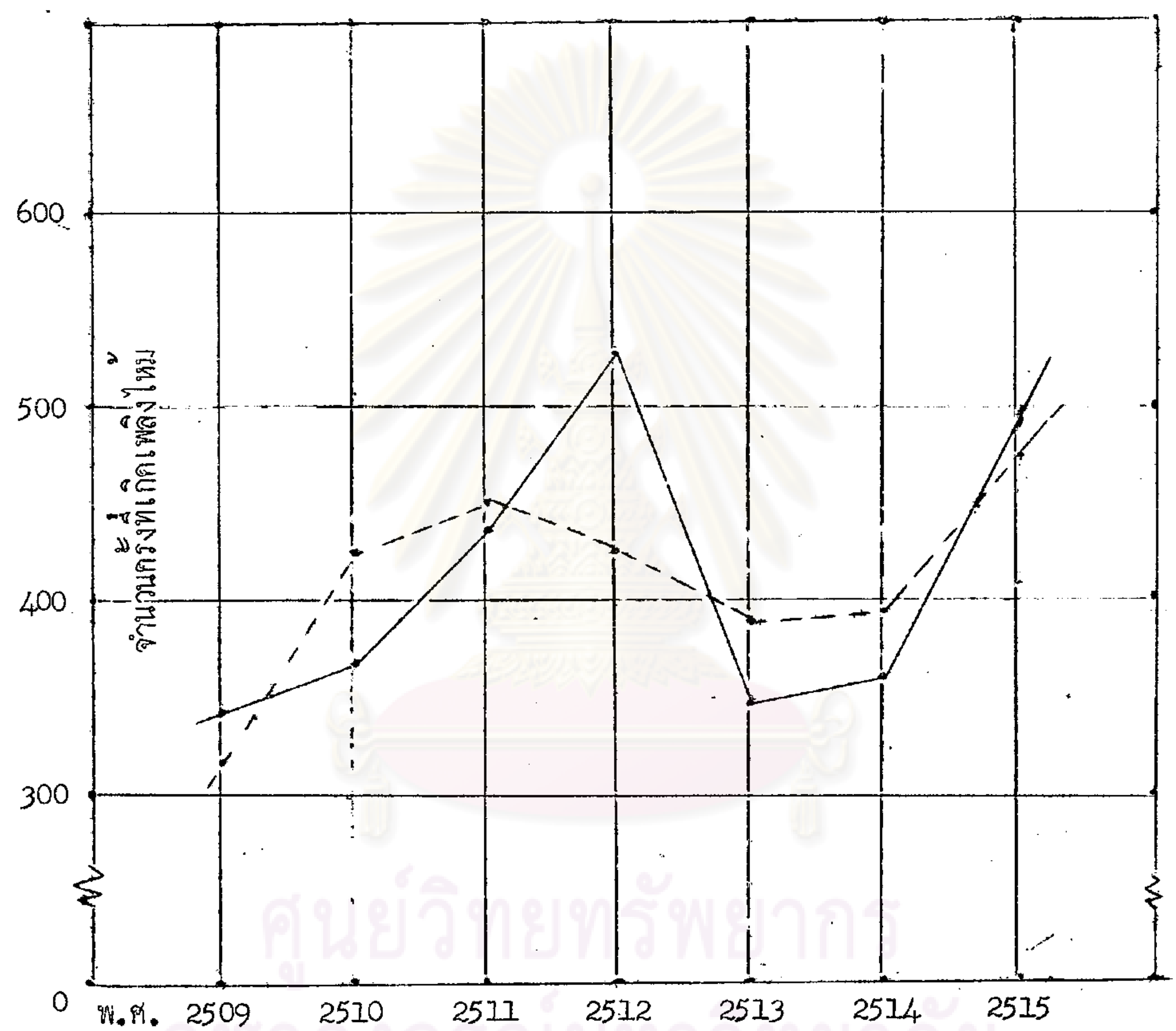
สาเหตุที่ใช้การทดสอบไคสแควร์ ทดสอบภาวะสารูปสนิทดีระหว่างโปลิโนเมียลโมเดล กับข้อมูลเดิม คือ ผู้เขียนวิทยานิพนธ์มีจุดมุ่งหมายสร้างโปลิโนเมียลโมเดลสำหรับประมาณข้อมูลเดิมว่าโปลิโนเมียลอันดับที่สามหรืออันดับที่สี่ หรืออันดับที่ห้าที่จะใช้แทนข้อมูลเดิมได้เหมาะสมกว่ากัน จึงจำเป็นต้องทดสอบภาวะสารูปสนิทดีระหว่างโปลิโนเมียลโมเดลดังกล่าว กับข้อมูลเดิมด้วยการทดสอบไคสแควร์

เมื่อสร้างโปลิโนเมียลโมเดลสำหรับแทนข้อมูลเกี่ยวกับจำนวนครั้งที่เกิดเพลิงไหม้ เป็นโปลิโนเมียลอันดับที่สาม สี่ และห้า ในทางปฏิบัติจะต้องตัดสินใจว่าโมเดลแบบใดที่จะเหมาะสมกับข้อมูลเดิม โดยการทดสอบข้อมูลเดิมกับโมเดลแต่ละแบบ สำหรับการทดสอบโมเดลจะใช้การทดสอบแบบไคสแควร์ โดยตั้งข้อสมมุติฐานว่า โปลิโนเมียลโมเดลที่สร้างขึ้นมาแต่ละแบบ ไม่มีความแตกต่างกับข้อมูลเดิม โดยใช้ระดับนัยสำคัญ 0.05 และ 0.01 ตามลำดับ

จากโมเดลโปลิโนเมียลอันดับที่สาม สี่ และห้า นำมาคำนวณหาจำนวนครั้งที่เกิดเพลิงไหม้เปรียบเทียบข้อมูลเดิม ดังตารางที่ 8 และภาพเปรียบเทียบที่ 5, 6 และ 7

ตารางที่ 8 แสดงการเปรียบเทียบจำนวนครั้งของการเกิดเพลิงไหม้จริง
กับโปลิโนเมียลอันดับที่สาม สี่ และห้า

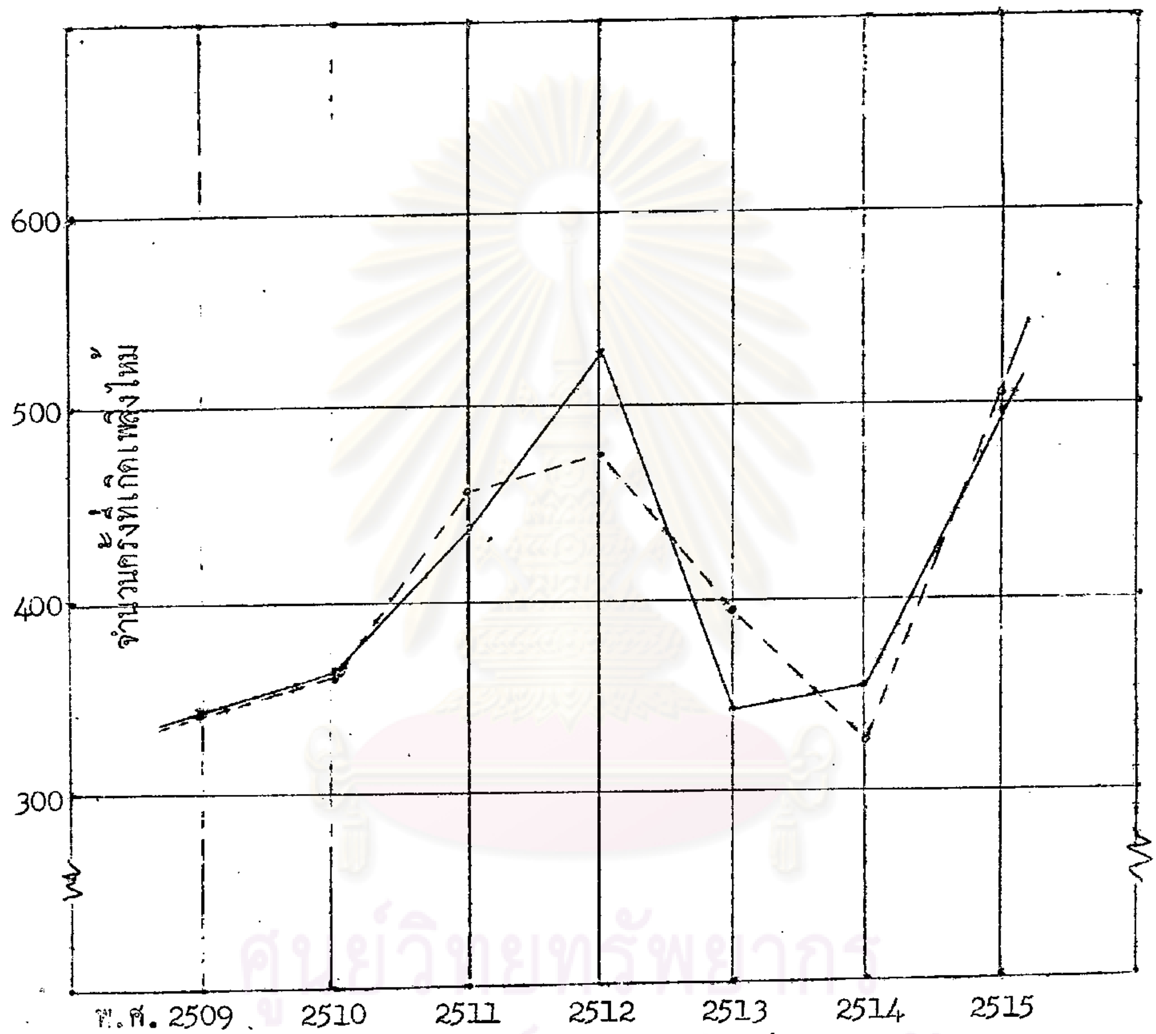
พ.ศ.	t	Observed Y_t	Polynomial order 3 Y_t	Polynomial order 4 Y_t	Polynomial order 5 Y_t
2509	1	343	314.547 619	342.034 531	345.558 440
2510	2	366	428.880 950	364.744 573	350.649 336
2511	3	439	450.595 229	459.757 497	477.376 578
2512	4	529	422.857 120	477.830 921	477.831 110
2513	5	342	388.833 287	397.995 067	380.376 684
2514	6	357	391.690 394	327.552 861	341.649 900
2515	7	496	474.595 105	502.079 933	498.560 246
	รวม	2872	2871.999 704	2871.995 483	2872.002 294



ภาพที่ 5 แสดงการเปรียบเทียบข้อมูลเดิมกับโปลิโนเมียลอันดับที่สาม

———— แทนข้อมูลเดิม

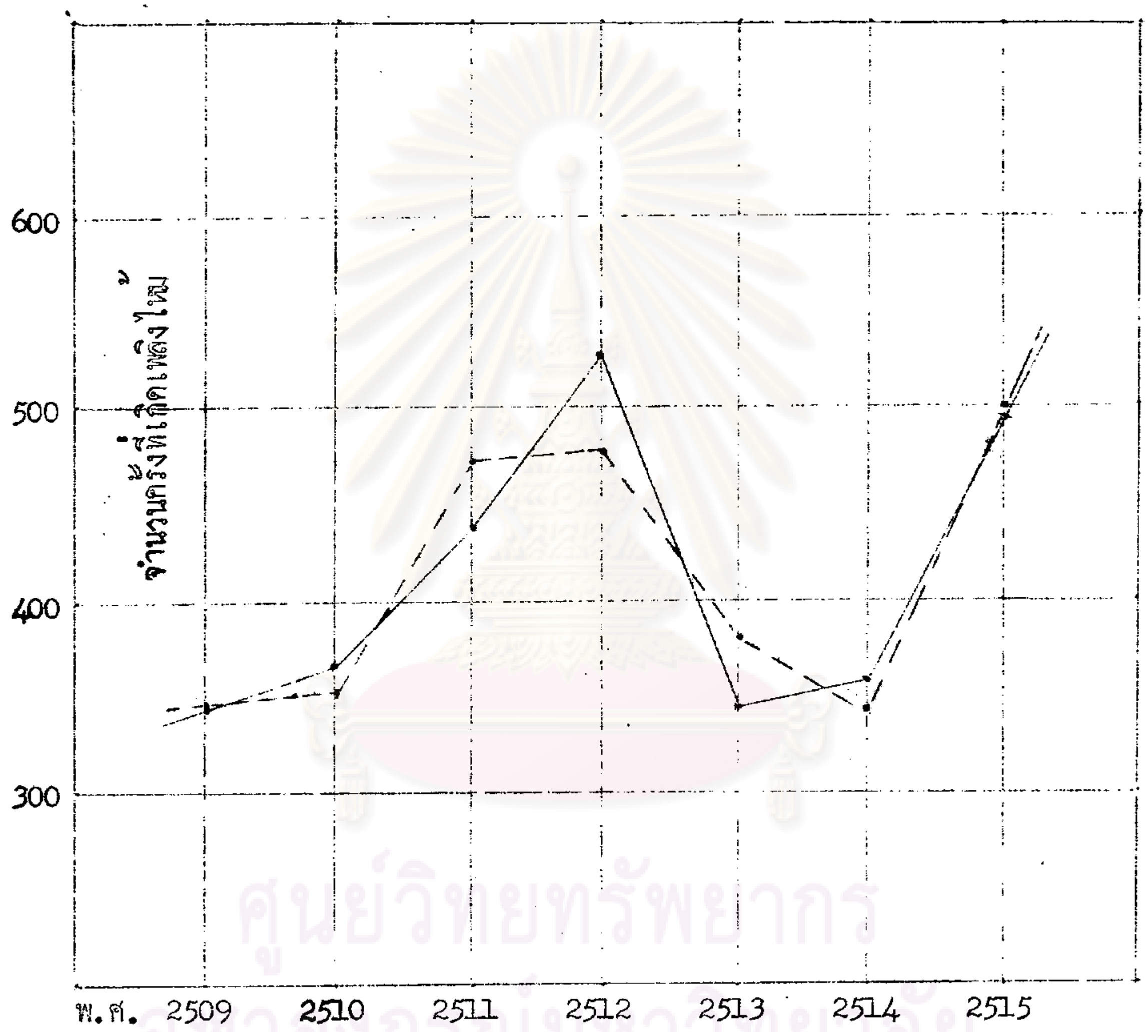
- - - - แทนโปลิโนเมียลอันดับที่สาม



ภาพที่ 6 แสดงการเปรียบเทียบข้อมูลเดิมกับโปลิโนเมียลอันดับที่สี่

——— แทนข้อมูลเดิม

- - - - แทนโปลิโนเมียลอันดับที่สี่



ภาพที่ 7 แสดงการเปรียบเทียบข้อมูลเดิมกับโปลิโนเมียลอันดับที่ห้า

——— แทนข้อมูลเดิม

– – – – แทนโปลิโนเมียลอันดับห้า

นำผลของการคำนวณแต่ละโมเดลจากตารางที่ 8 ทดสอบกับข้อมูลเดิม โดยใช้ไคสแควร์ทดสอบ โดยกำหนดให้

f_o = จำนวนครั้งที่เกิดเพลิงไหม้ในแต่ละปี

f_e = จำนวนครั้งที่เกิดเพลิงไหม้ที่คำนวณได้จากโปลิโนเมียล

d.f. = (7 - 1) (2 - 1) = 6

ผลของการทดสอบโปลิโนเมียลอันดับที่สาม สี่และห้า แสดงอยู่ในตารางที่ 9, 10 และ 11 ตามลำดับ

จากการเปิดตารางไคสแควร์ Degree of freedom 6

ณ ระดับนัยสำคัญ 0.05 d.f. = 6 χ^2 = 12.59

ณ ระดับนัยสำคัญ 0.01 d.f. = 6 χ^2 = 16.81

ผลของการทดสอบสรุปได้ดังนี้

1. ณ ระดับนัยสำคัญ 0.05 d.f. = 6 โปลิโนเมียลอันดับที่สาม สี่ และห้า มีนัยสำคัญคือ ไม่ยอมรับสมมุติฐานที่ว่า โมเดลที่สร้างขึ้นมากับข้อมูลเดิมไม่มีความแตกต่างกัน นั่นก็คือ โมเดลที่สร้างขึ้นมาไม่มีภาวะสารูปสนิทดี (Goodness of Fit) กับข้อมูลเดิม ณ ระดับนัยสำคัญดังกล่าว

2. ณ ระดับนัยสำคัญ 0.01 d.f. = 6 โปลิโนเมียลอันดับที่สาม และสี่ มีนัยสำคัญเช่นเดียวกับข้อสรุป 1

3. ณ ระดับนัยสำคัญ 0.01 d.f = 6 โปลิโนเมียลอันดับที่ห้า ไร้นัยสำคัญคือ ยอมรับสมมุติฐานที่ว่า โมเดลที่สร้างขึ้นมามีภาวะสารูปสนิทดี (Goodness of Fit) กับข้อมูลเดิม ซึ่งเป็นข้อสนับสนุนที่ว่า โปลิโนเมียลอันดับที่ห้าเหมาะสมที่จะเป็นโมเดลแทนข้อมูลเดิม

ตารางที่ 9 แสดงการทดสอบ χ^2 ของข้อมูลเดิมกับโปลิโนเมียลอันดับที่สาม

ข้อมูลเดิม f_o	ข้อมูลจากการคำนวณ f_e	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
343	314.548	+ 28.452	809.516	2.5735
366	428.881	- 62.881	3954.020	9.2193
439	450.595	- 11.595	134.444	0.2983
529	422.857	+ 106.143	11266.336	26.6430
342	388.833	- 46.833	2193.330	5.6408
357	391.690	- 34.690	1203.396	3.0723
496	474.595	+ 21.405	458.174	0.9653
2872	2871.999		χ_c^2 =	48.4127

จากตารางไคสแควร์ เมื่อ d.f = 6

ณ ระดับนัยสำคัญ 0.05 $\chi^2 = 12.59 < \chi_c^2$

ณ ระดับนัยสำคัญ 0.01 $\chi^2 = 16.81 < \chi_c^2$

ตารางที่ 10 แสดงการทดสอบไคสแควร์ข้อมูลเดิมกับโปลิโนเมียลอันดับที่สี่

f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
343	342.035	+ 0.965	0.931	0.0027
366	364.745	+ 1.255	1.575	0.0043
439	459.757	- 20.757	430.895	0.9372
529	477.831	+ 51.169	2618.267	5.4794
342	397.995	- 55.995	3135.440	7.8780
357	327.553	+ 29.447	867.126	2.6472
496	502.080	- 6.080	36.966	0.0736
2872	2871.996		χ^2_c =	17.0224

จากตารางไคสแควร์ เมื่อ d.f = 6

ณ ระดับนัยสำคัญ 0.05 $\chi^2 = 12.59 < \chi^2_c$

ณ ระดับนัยสำคัญ 0.01 $\chi^2 = 16.81 < \chi^2_c$

ตารางที่ 11 แสดงการทดสอบไคสแควร์ข้อมูลเดิมกับโปลิโนเมียลอันดับที่ห้า

f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
343	345.558	- 2.558	6.543	0.0189
366	350.649	+ 15.351	235.653	0.6720
439	477.376	- 38.376	1472.794	3.0851
529	477.831	+ 51.169	2618.266	5.4794
342	380.376	- 38.376	1472.794	3.8719
357	341.649	+ 15.351	235.622	0.6896
496	498.560	- 2.560	6.553	0.0131
2872	2872.002		χ^2_c =	13.8300

จากตารางไคสแควร์ เมื่อ d.f = 6

ณ ระดับนัยสำคัญ 0.05 $\chi^2 = 12.59 < \chi^2_c$

ณ ระดับนัยสำคัญ 0.01 $\chi^2 = 16.81 > \chi^2_c$