บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

ผลการวิจัยครั้งนี้จะเป็นประโยชน์แก่ผู้วิจัยทั่วไปที่ต้องการรูปแบบการประมาณการแจกแจงของประชากรที่มีการแจกแจงปัวส์ซง จากการวิจัยทำให้ทราบว่า วิธีการประมาณฟังก์ชันความหนาแน่นของประชากรด้วยวิธีเคอร์เนล ซึ่งเป็นวิธีที่เหมาะสมกับตัวแปรสุ่มต่อเนื่อง ใช้ได้ดีสำหรับการประมาณการแจกแจงปัวส์ซงซึ่งเป็นการแจกแจงของตัวแปรสุ่มไม่ต่อเนื่องด้วย ข้อมูลที่ใช้ในการศึกษาได้จากการจำลองขึ้นมาโดยใช้เครื่องคอมพิวเตอร์ ขนาดของกลุ่มตัวอย่าง (n) ที่ศึกษาคือ 30 50 และ 100 ข้อมูลจำลองขึ้นมาแต่ละครั้งต้องทำการทดสอบภาวะสารูปดีที่ $\alpha = 0.05$ เพื่อตรวจสอบว่ามีการแจกแจงปัวส์ซงพารามิเตอร์ λ เป็น 1 2 3 4 และ 5 หรือไม่ สำหรับขนาดของกลุ่มตัวอย่างแต่ละขนาดและ λ แต่ละค่าจะทำการจำลองข้อมูลซ้ำ ๆ กัน 200 ครั้ง และใช้ข้อมูลที่จำลองได้ศึกษาถึงค่าคาดหวังของค่าฟังก์ชัน $\hat{f}$ ที่ประมาณได้ ค่าผลรวมความคลาดเคลื่อนกำลังสองเฉลี่ยของ $\hat{f}$ ค่าเอียงเฉยกกำลังสองของ $\hat{f}$ และค่าความแปรปรวนของ $\hat{f}$ โดยใช้วิธีการประมาณด้วยวิธีเคอร์เนล ฟังก์ชันเคอร์เนลที่ศึกษาคือ K_1 และ K_2 ส่วนความกว้าง ศึกษาความกว้าง h_1 h_2 และ h_3 พร้อมทั้งประมาณค่า $\hat{f}(x)$ เมื่อ $x = 0, 1, 2, \ldots$ ด้วยวิธีหาความถี่สัมพัทธ์ ประกอบการพิจารณาด้วย จากวิธีการประมาณดังกล่าว เลือกวิธีประมาณฟังก์ชัน $\hat{f}$ ที่ทำให้ค่าผลรวมความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด ผลการวิจัยสรุปได้ดังนี้

ตารางที่ 5.1 แสดงวิธีการประมาณฟังก์ชัน $\hat{f}$ ที่ทำให้ค่าผลรวมความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำที่สุด

n \ λ	1	2	3	4	5
30	K_2h_3	K_2h_1	K_2h_1	K_2h_1	K_2h_1
50	K_2h_3	K_2h_1	K_2h_1	K_2h_1	K_2h_1
100	K_2h_3	K_2h_1	K_2h_1	K_2h_1	K_2h_1

จากตารางที่ 5.1 สรุปได้ว่า วิธีการประมาณฟังก์ชันความน่าจะเป็นของการแจกแจงปัวส์ซงที่มีพารามิเตอร์ λ เป็น 1 2 3 4 และ 5 วิธีที่เหมาะสมคือ วิธีเคอร์เนล ส่วนความกว้างใช้แตกต่างกันตามค่าพารามิเตอร์ คือ $\lambda = 1$ ใช้ความกว้าง $h_3 = 0.5$ ถ้า $\lambda = 2$ 3 4 5 ใช้ความกว้าง $h_1 = 1.06\sigma n^{-(1/5)}$

ตารางที่ 5.2 แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 1$

$n = 30 \quad 50$ และ 100

n	30	50	100	
x	$\hat{f}(x)$	$\hat{f}(x)$	$\hat{f}(x)$	$f(x)$
0	.3337	.3335	.3348	.36788
1	.3570	.3573	.3563	.36788
2	.1910	.1910	.1897	.18394
3	.0682	.0691	.0699	.06131
4	.0195	.0187	.0190	.01533
5	.0044	.0042	.0040	.00307
6	.0006	.0007	.0005	.00051
7	.0000	.0000	.0001	.00007
ผลรวม	.9745	.9745	.9743	.99999

ตารางที่ 5.3 แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 2$

$n = 30$ 50 และ 100

n	30	50	100	
x	$\hat{f}(x)$	$\hat{f}(x)$	$\hat{f}(x)$	$f(x)$
0	.1349	.1361	.1349	.13534
1	.2403	.2476	.2531	.27067
2	.2458	.2500	.2586	.27067
3	.1771	.1768	.1801	.18045
4	.0991	.0986	.0960	.09022
5	.0443	.0426	.0407	.03609
6	.0164	.0143	.0136	.01203
7	.0046	.0036	.0036	.00344
8	.0011	.0010	.0007	.00086
9	.0004	.0003	.0001	.00019
10	.0001	.0000	.0001	.00004
ผลรวม	.9640	.9709	.9814	.99999

ตารางที่ 5.4 แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 3$

$n = 30$ 50 และ 100

n	30	50	100	
x	$\hat{f}(x)$	$\hat{f}(x)$	$\hat{f}(x)$	$f(x)$
0	.0639	.0641	.0625	.04979
1	.1389	.1424	.1440	.14936
2	.2003	.2037	.2073	.22404
3	.2078	.2089	.2113	.22404
4	.1646	.1640	.1648	.16803
5	.1066	.1041	.1035	.10082
6	.0561	.0554	.0542	.05041
7	.0254	.0244	.0241	.02160
8	.0104	.0100	.0095	.00810
9	.0039	.0039	.0034	.00270
10	.0013	.0013	.0012	.00081
11	.0004	.0004	.0005	.00022
12	.0001	.0002	.0002	.00006
13	.0000	.0001	.0000	.00001
ผลรวม	.9798	.9829	.9865	1.00000

ตารางที่ 5.5 แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 4$

n = 30 50 และ 100

n	30	50	100	
x	$\hat{f}(x)$	$\hat{f}(x)$	$\hat{f}(x)$	$f(x)$
0	.0354	.0324	.0300	.01832
1	.0809	.0794	.0784	.07326
2	.1358	.1373	.1397	.14653
3	.1745	.1779	.1821	.19537
4	.1776	.1813	.1835	.19537
5	.1487	.1498	.1507	.15629
6	.1047	.1049	.1050	.10420
7	.0639	.0645	.0632	.05954
8	.0348	.0347	.0334	.02977
9	.0175	.0168	.0162	.01323
10	.0081	.0072	.0071	.00529
11	.0033	.0025	.0025	.00192
12	.0011	.0008	.0006	.00064
13	.0004	.0003	.0001	.00020
14	.0001	.0002	.0001	.00006
15	.0001	.0001	.0000	.00002
ผลรวม	.9867	.9900	.9927	1.00000

ตารางที่ 5.6 แสดงค่าของ $\hat{f}$ ที่ประมาณได้ และค่าของ f แท้จริง เมื่อ $\lambda = 5$

$n = 30 \quad 50$ และ 100

n	30	50	100	
x	$\hat{f}(x)$	$\hat{f}(x)$	$\hat{f}(x)$	$f(x)$
0	.0199	.0176	.0150	.00674
1	.0472	.0451	.0422	.03369
2	.0873	.0869	.0860	.08422
3	.1292	.1312	.1336	.14037
4	.1561	.1599	.1637	.17547
5	.1567	.1610	.1651	.17547
6	.1371	.1378	.1408	.14622
7	.1040	.1036	.1044	.10444
8	.0695	.0693	.0683	.06528
9	.0411	.0410	.0396	.03627
10	.0221	.0216	.0205	.01813
11	.0113	.0105	.0097	.00824
12	.0054	.0047	.0041	.00343
13	.0024	.0020	.0018	.00132
14	.0010	.0009	.0008	.00047
15	.0004	.0005	.0003	.00016
16	.0002	.0002	.0001	.00005
17	.0001	.0001	.0000	.00001
ผลรวม	.9919	.9940	.9961	.99999

จากตารางที่ 5.2 5.3 5.4 5.5 และ 5.6 สรุปได้ว่า ขนาดของกลุ่มตัวอย่างอาจมีผลต่อค่าของ $\hat{f}$ ที่ประมาณได้บ้าง แต่ไม่มากนัก โดยขนาดของกลุ่มตัวอย่างที่ใหญ่ขึ้นจะทำให้ผลรวมของค่า $\hat{f}$ ที่ประมาณได้ใกล้เคียง 1 มากขึ้นด้วย

นอกจากนี้ผู้วิจัยยังพบจากการศึกษาเพิ่มเติมว่า ถ้าประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ $\lambda = 0.75$ วิธีการประมาณฟังก์ชันความน่าจะเป็นเหมือนกับเมื่อ $\lambda = 1$ คือใช้วิธีเคอร์เนล มีฟังก์ชันความหนาแน่นปกติมาตรฐาน (K_2) เป็นฟังก์ชันเคอร์เนล และมี $h_3 = 0.5$ เป็นความกว้างเช่นเดียวกัน แต่ถ้า $\lambda = 0.5$ ใช้วิธีหาความถี่สัมพัทธ์เป็นวิธีที่เหมาะสมที่สุด ดังนั้นผู้วิจัยจึงใช้การประมาณฟังก์ชันความน่าจะเป็นของจำนวนเซลล์ที่มีสารอัลคาลอยด์ในใบหญ้าเกล็ดปลาด้วยวิธีหาความถี่สัมพัทธ์เมื่อ $\hat{\lambda} = 0.5$ และใช้วิธีเคอร์เนล มีฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนล และความกว้าง $h_3 = 0.5$ เมื่อ $\hat{\lambda} = 0.74$ และ 0.76

5.2 ข้อเสนอแนะ

5.2.1 ผลการศึกษาพบว่า วิธีการประมาณการแจกแจงปัวส์ซงด้วยวิธีเคอร์เนลเป็นวิธีการประมาณที่ให้ค่าผลรวมความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด แต่มีปัญหาเกิดขึ้นคือ ผลรวมของค่า $\hat{f}$ ซึ่งเป็นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มปัวส์ซง มักมีค่าไม่เป็น 1 ทั้งนี้อาจเนื่องจากสาเหตุหลายประการ คือ

ก. เนื่องจากใช้วิธีการประมาณฟังก์ชันความหนาแน่นของตัวแปรสุ่มต่อเนื่องไปใช้กับตัวแปรสุ่มไม่ต่อเนื่อง ในการหาผลรวม $\hat{f}$ นั้นเท่ากับเป็นการใช้ผลรวมไปแทนการอินทิเกรท

ข. อาจมีค่าของ $\hat{f}(x) > 0$ ที่จุด x บางจุด เช่น เมื่อ x น้อยกว่า 0 หรือ x มีค่ามาก ๆ ซึ่งไม่ได้ถูกรวมค่าเข้าไป

จากการวิจัยพบว่า ที่ค่า λ เดียวกัน ขนาดของกลุ่มตัวอย่างที่ใหญ่ขึ้น มีผลทำให้ผลรวมของค่า $\hat{f}$ ใกล้ 1 มากขึ้น ดังนั้นการใช้กลุ่มตัวอย่างขนาดใหญ่ จะช่วยลดความคลาดเคลื่อนลงได้

5.2.2 ความกว้าง ควรลองศึกษาความกว้างอื่น ๆ เพิ่มเติมจากการวิจัยครั้งนี้ โดยผู้วิจัยเลือกกำหนดให้เหมาะสมกับลักษณะพื้นฐานของการแจกแจงของประชากรที่นำมาศึกษา

5.2.3 กรณีที่ขนาดของกลุ่มตัวอย่างใหญ่ขึ้น ทำให้ฟังก์ชันความน่าจะเป็นที่ประมาณได้มีความถูกต้องมากขึ้น กล่าวคือ ผลรวมของค่า $\hat{f}$ ใกล้เคียง 1 มากขึ้น และค่าผลรวมความคลาดเคลื่อนกำลังสองเฉลี่ยลดลงนั้น น่าจะมีการศึกษาเพิ่มเติมว่ามาจากสาเหตุอะไร