

บทที่ 4

ผลการวิจัย

ผู้วิจัยทำการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย (*MISE*) และค่าแบนวิตซ์เฉลี่ย ที่เกิดขึ้นจากการประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นทั้ง 2 รูปแบบ โดยใช้ค่าแบนวิตซ์จากวิธีการเลือกแบนวิตซ์ทั้ง 6 วิธี ในการวิเคราะห์ผลจะใช้เกณฑ์ในการเปรียบเทียบ คือ ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยเป็นตัวชี้วัด ผลการวิจัยจำแนกได้ 2 ลักษณะ คือ ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย และ ค่าแบนวิตซ์เฉลี่ย นำเสนอในรูปแบบตารางและรูปภาพกราฟ รวมทั้งแสดงตัวอย่างกราฟความหนาแน่นที่ได้จากการประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็น จำแนกตามรูปแบบฟังก์ชันความหนาแน่นของความน่าจะเป็น วิธีการเลือกแบนวิตซ์ และขนาดตัวอย่าง การวิจัยดังกล่าวเป็นการศึกษาเชิงจำลอง (Simulation study) แสดงผลการวิจัยได้ในหัวข้อที่ 4.1

นอกจากการศึกษาเชิงจำลองแล้ว ผู้วิจัยได้ทำการศึกษาข้อมูลเชิงตัวเลข (Numerical data study) โดยนำการประมาณความหนาแน่นด้วยวิธีของ Zhang และคณะ (1999) มาประยุกต์ใช้กับข้อมูลจริง ข้อมูลที่นำมาวิจัยจะเป็นชุดข้อมูลที่มีลักษณะกราฟฟังก์ชันความหนาแน่นของความน่าจะเป็นที่เกิดปัญหาการประมาณความหนาแน่นบริเวณขอบเขต (ปัญหา Boundary effect) ผลการวิจัยของการศึกษาข้อมูลเชิงตัวเลขแสดงได้ในหัวข้อที่ 4.2

สัญลักษณ์ต่าง ๆ ที่ใช้ในการนำเสนอมีรายละเอียดดังนี้

DPI	แทน	การเลือกแบนวิตซ์ด้วยวิธี	Direct plug-in
ROT	แทน	การเลือกแบนวิตซ์ด้วยวิธี	Rules of thumb
LSCV	แทน	การเลือกแบนวิตซ์ด้วยวิธี	Least squares cross-validation
SROT	แทน	การเลือกแบนวิตซ์ด้วยวิธี	Silverman's rules of thumb
BCV	แทน	การเลือกแบนวิตซ์ด้วยวิธี	Biased cross-validation
STE	แทน	การเลือกแบนวิตซ์ด้วยวิธี	Solve the equation plug-in
MISE	แทน	ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย	
n	แทน	ขนาดตัวอย่าง	

4.1 ผลการวิจัยของการศึกษาเชิงจำลอง

ผู้วิจัยได้วางแผนการทดลองของการศึกษาเชิงจำลอง ภายใต้สถานการณ์ที่กำหนด
ดังนี้

กำหนดฟังก์ชันเคอร์เนลแบบเกาส์เซียน ใช้ขนาดตัวอย่าง (n) จำนวน 5 ขนาด คือ 30 , 50 , 100 , 150 และ 200 ค่าแบนวิตซ์เกิดจากวิธีการเลือกแบนวิตซ์จำนวน 6 วิธี คือ วิธี Direct plug-in วิธี Rules of thumb วิธี Least squares cross-validation วิธี Silverman's rules of thumb วิธี Biased Cross-Validation และ วิธี Solve the equation plug-in ตัวแปรสุ่มที่ใช้ในการวิจัยเกิดจากฟังก์ชันความหนาแน่นของความน่าจะเป็น 2 รูปแบบ จำแนกตามกรณีศึกษาได้ ดังนี้

1. กรณีศึกษาที่ 1 ตัวแปรสุ่มเกิดจากฟังก์ชันความหนาแน่นของความน่าจะเป็น
รูปแบบที่ 1 คือ $f(x) = 5e^{-5x}$

2. กรณีศึกษาที่ 2 ตัวแปรสุ่มเกิดจากฟังก์ชันความหนาแน่นของความน่าจะเป็น
รูปแบบที่ 2 คือ $f(x) = \frac{5}{4}(1 + 15x)e^{-5x}$

4.1.1 กรณีศึกษาที่ 1

ผลการเปรียบเทียบประสิทธิภาพของวิธีการเลือกแบนวิดท์ แสดงได้ดังตารางที่ 1

ตารางที่ 1

ผลการเปรียบเทียบประสิทธิภาพของวิธีการเลือกแบนวิดท์ 6 วิธี กรณีตัวแปรสุ่ม
เกิดจากฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1

ขนาดตัวอย่าง (n)	วิธีการเลือกแบนวิดท์	MISE	แบนวิดท์เฉลี่ย
30	DPI	0.2009453	0.06012773
	ROT	0.1356997*	0.1039066
	LSCV	0.3503985**	0.04978907
	SROT	0.1765491	0.06999994
	BCV	0.1743319	0.1264269
	STE	0.2469504	0.05153638
50	DPI	0.1205716	0.05268311
	ROT	0.07805198*	0.09463914
	LSCV	0.2421534**	0.03784206
	SROT	0.09388683	0.06490549
	BCV	0.1030709	0.1015518
	STE	0.1565082	0.04410024
100	DPI	0.06531306	0.04243371
	ROT	0.03873861*	0.08357637
	LSCV	0.1454323**	0.02514043
	SROT	0.04291475	0.05796355
	BCV	0.04736212	0.07171911
	STE	0.08376166	0.03514478

* ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยน้อยที่สุด

** ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมากที่สุด

ตารางที่ 1 (ต่อ)

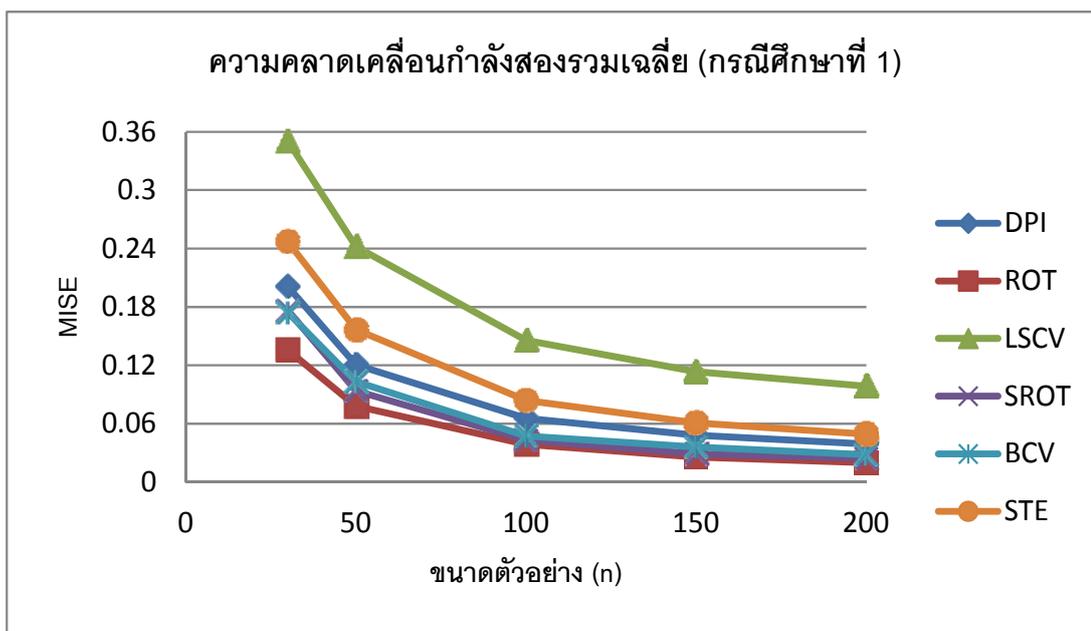
ขนาดตัวอย่าง (n)	วิธีการเลือกแบนด์วิดท์	MISE	แบนด์วิดท์เฉลี่ย
150	DPI	0.04801244	0.03739146
	ROT	0.02583383*	0.07761845
	LSCV	0.1133199**	0.02022155
	SROT	0.02872980	0.05336375
	BCV	0.03587144	0.05608545
	STE	0.0605862	0.03070374
200	DPI	0.03884649	0.03384478
	ROT	0.01981895*	0.07251839
	LSCV	0.09826**	0.01697422
	SROT	0.02437209	0.0505765
	BCV	0.0278454	0.04842149
	STE	0.04943376	0.02762671

* ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยน้อยที่สุด

** ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมากที่สุด

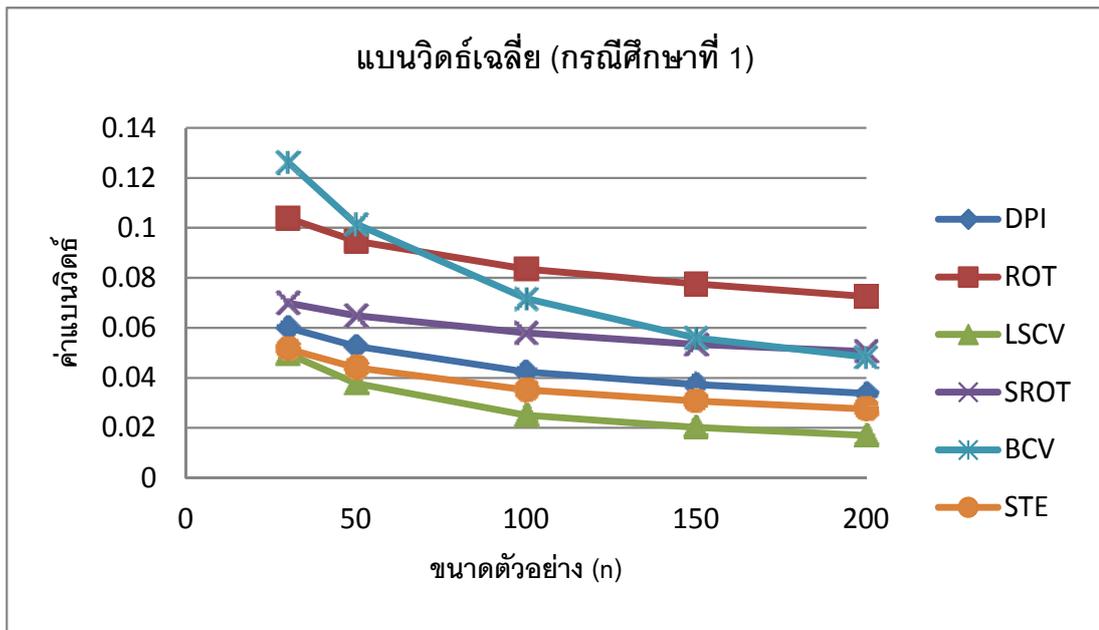
ภาพที่ 4.1

การเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย กรณีตัวแปรสุ่ม
เกิดจากฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1



จากตารางที่ 1 และภาพที่ 4.1 การเลือกแบนวิดธ์ด้วยวิธี Rules of thumb จะให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นน้อยที่สุดในทุกขนาดตัวอย่าง การเลือกแบนวิดธ์ด้วยวิธี Silverman's rules of thumb และ วิธี Biased cross-validation ให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นที่ใกล้เคียงกันในทุกขนาดตัวอย่างและให้ความคลาดเคลื่อนกำลังสองรวมเฉลี่ยที่ตรงจากวิธี Rules of thumb การเลือกแบนวิดธ์ด้วยวิธี Least squares cross-validation ให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นสูงที่สุดในทุกขนาดตัวอย่าง และค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

ภาพที่ 4.2
การเปรียบเทียบค่าแบนวิดธ์เฉลี่ย กรณีตัวแปรสุ่มเกิดจาก
ฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1



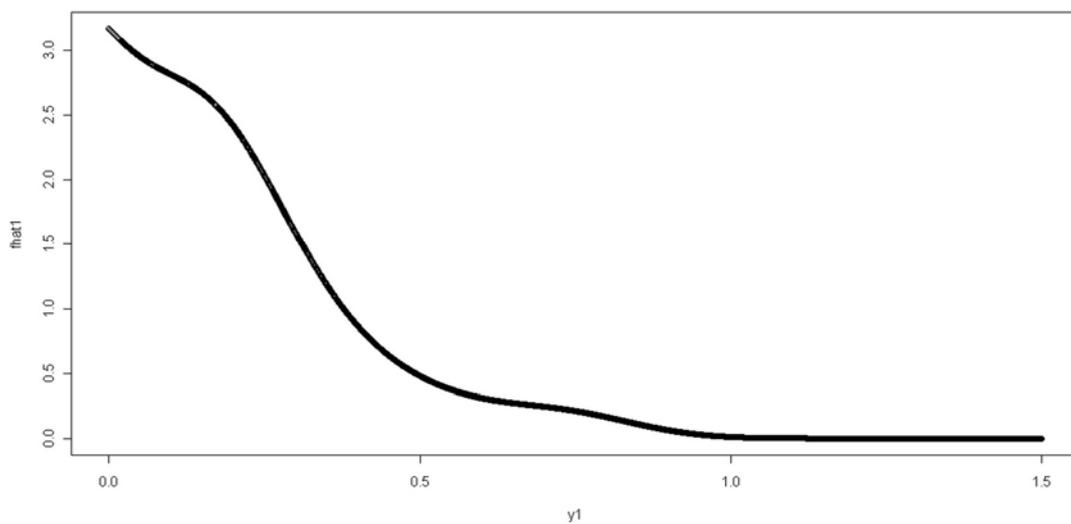
จากตารางที่ 1 และภาพที่ 4.2 วิธี Least squares cross-validation จะให้ค่าแบนวิดธ์เฉลี่ยที่น้อยที่สุดในทุกขนาดตัวอย่าง รองลงมาคือวิธี Solve the equation plug-in เมื่อขนาดตัวอย่างเท่ากับ 30 และ 50 วิธี Biased cross-validation ให้ค่าแบนวิดธ์เฉลี่ยมากที่สุด เมื่อขนาดตัวอย่างเท่ากับ 100 , 150 และ 200 วิธี Rules of thumb ให้ค่าแบนวิดธ์เฉลี่ยมากที่สุด และค่าแบนวิดธ์เฉลี่ยมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

ตัวอย่างกราฟความหนาแน่นที่ได้จากการประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 $f(x) = 5e^{-5x}$ โดยวิธีการเลือกแบนวิดธ์ที่ทำให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมีค่าน้อยที่สุดในทุกขนาดตัวอย่างคือวิธี Rules of thumb และค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมีค่ามากที่สุดในทุกขนาดตัวอย่างคือวิธี Least squares cross-validation แสดงได้ดังต่อไปนี้

กราฟความหนาแน่นที่ได้จากการประมาณความหนาแน่นโดยใช้ค่าแบนวิดท์ที่ได้จากวิธี Rules of thumb จำแนกตามขนาดตัวอย่าง แสดงได้ดังภาพที่ 4.3 ถึง ภาพที่ 4.7

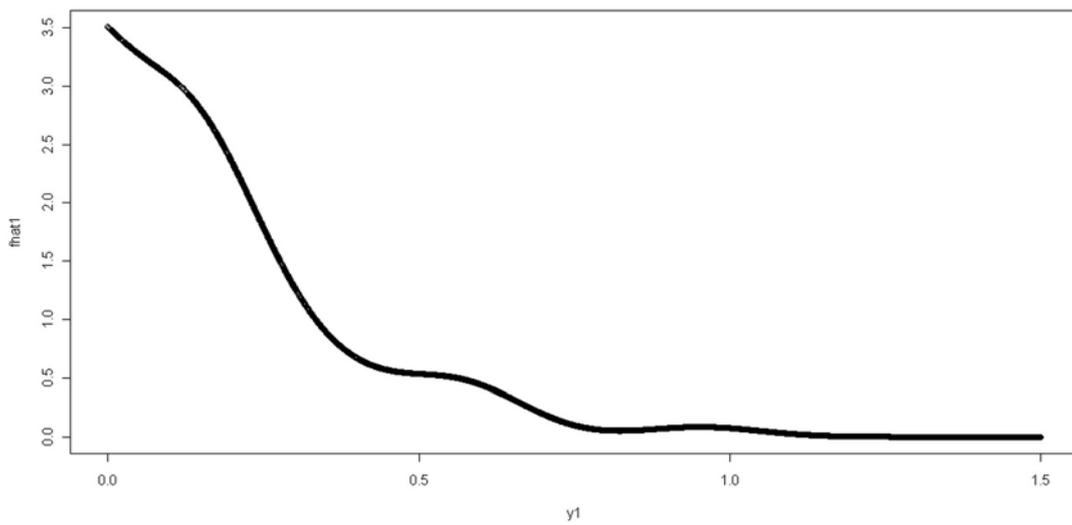
ภาพที่ 4.3

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิดท์จากการเลือกแบนวิดท์ด้วยวิธี Rules of thumb เมื่อขนาดตัวอย่างคือ 30



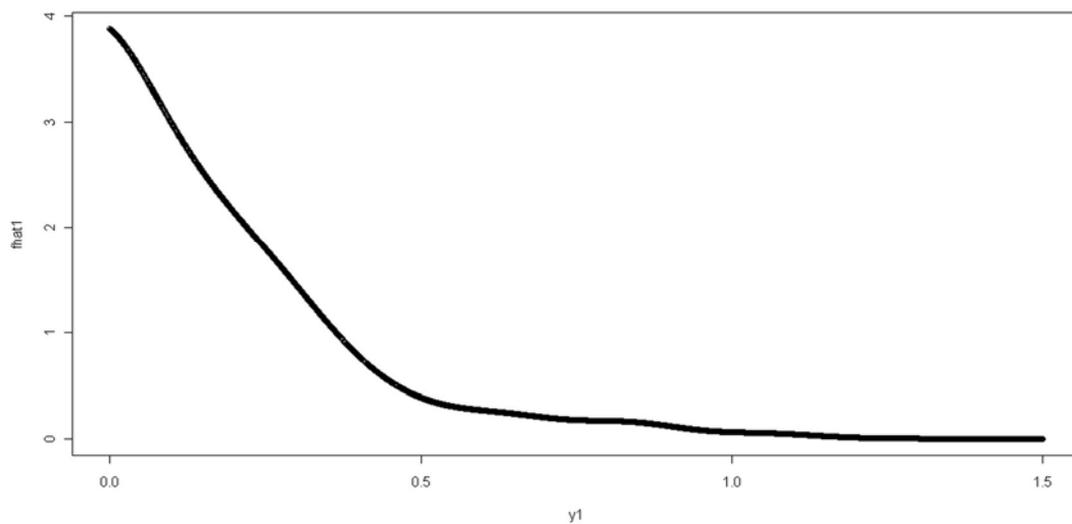
ภาพที่ 4.4

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิตซ์จากการเลือก
แบนวิตซ์ด้วยวิธี Rules of thumb เมื่อขนาดตัวอย่างคือ 50



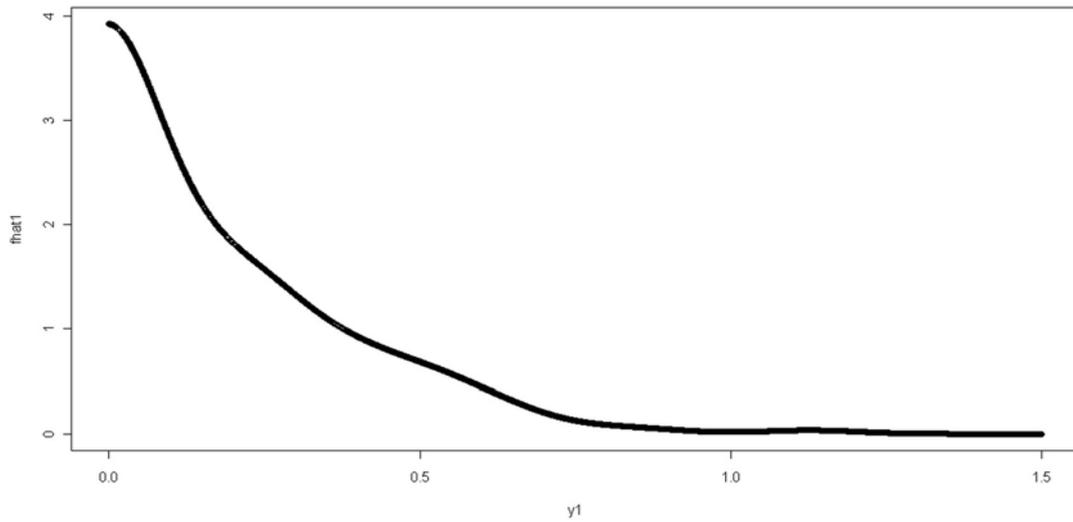
ภาพที่ 4.5

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิตซ์จากการเลือก
แบนวิตซ์ด้วยวิธี Rules of thumb เมื่อขนาดตัวอย่างคือ 100



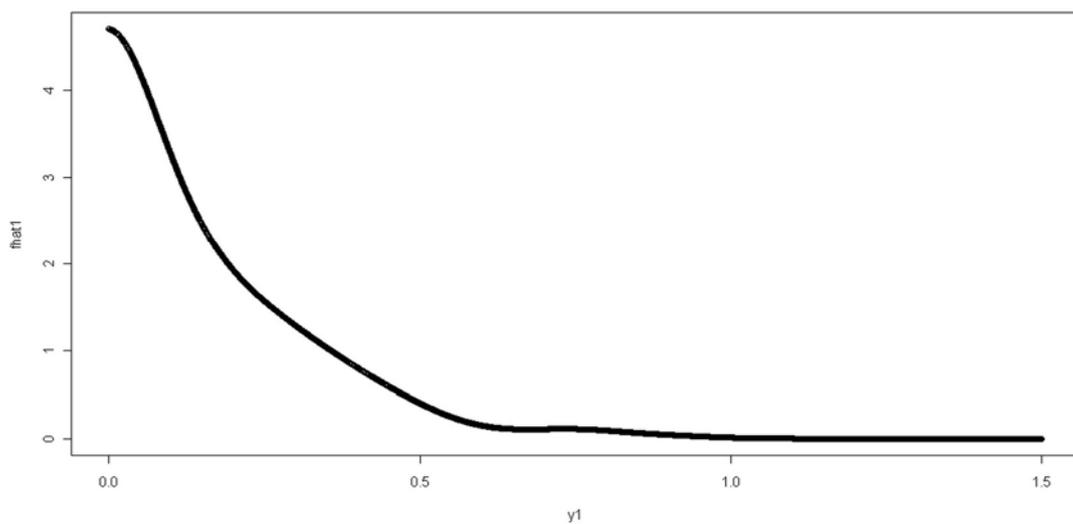
ภาพที่ 4.6

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิดธ์จากการเลือกแบนวิดธ์ด้วยวิธี Rules of thumb เมื่อขนาดตัวอย่างคือ 150



ภาพที่ 4.7

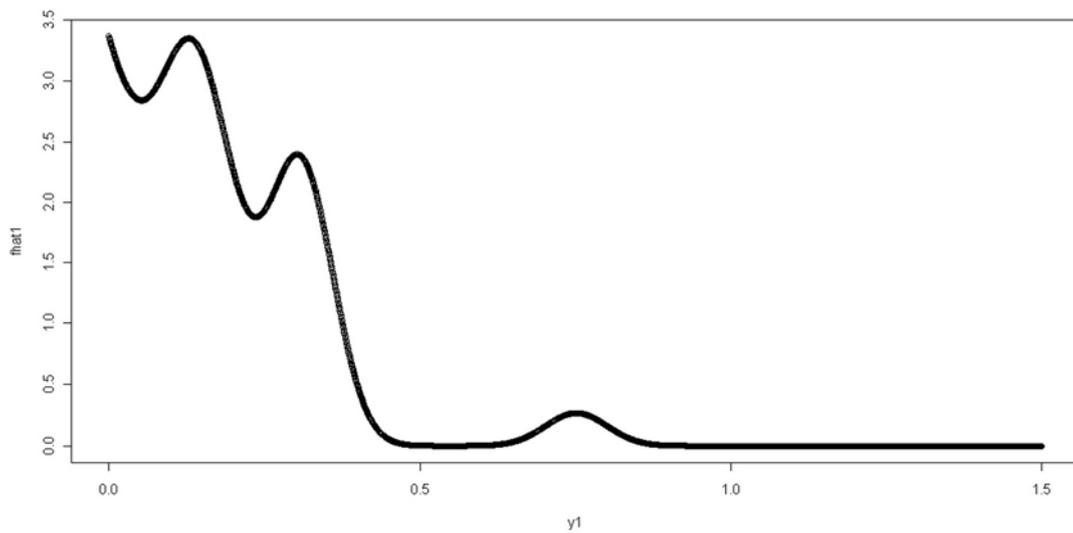
การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิดธ์จากการเลือกแบนวิดธ์ด้วยวิธี Rules of thumb เมื่อขนาดตัวอย่างคือ 200



กราฟความหนาแน่นที่ได้จากการประมาณความหนาแน่นโดยใช้ค่าแบนวิตซ์ที่ได้จากวิธี Least squares cross-validation จำแนกตามขนาดตัวอย่าง แสดงได้ดังภาพที่ 4.8 ถึง ภาพที่ 4.12

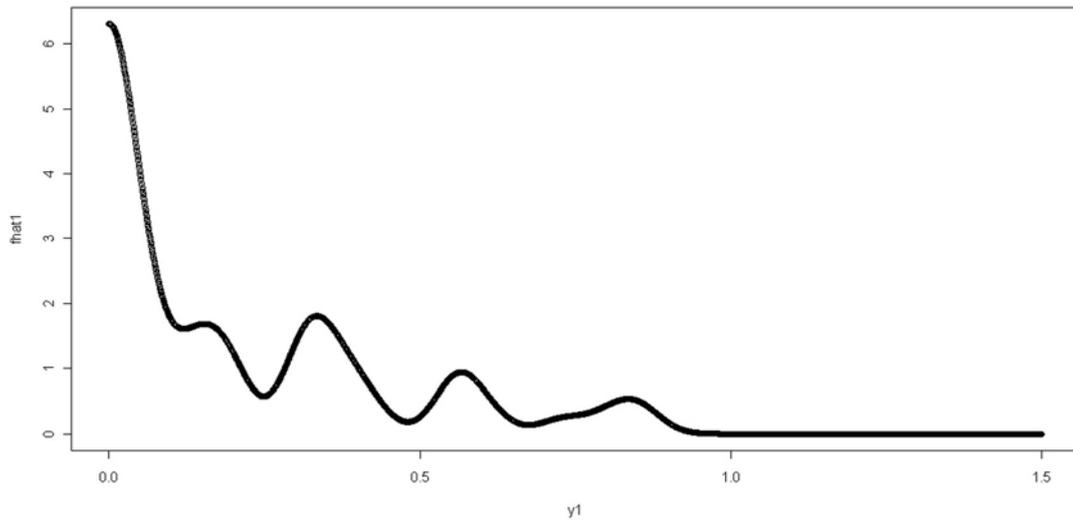
ภาพที่ 4.8

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 30



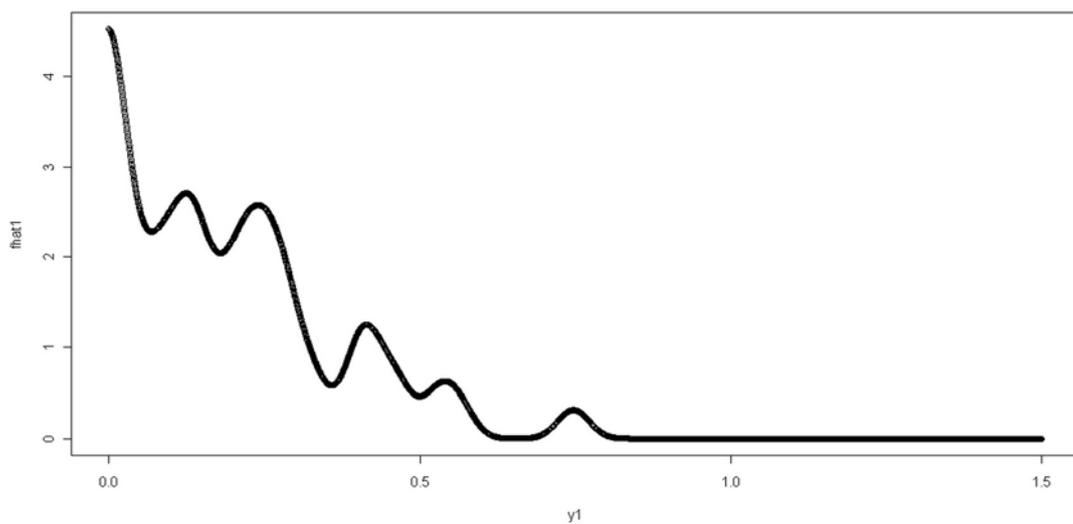
ภาพที่ 4.9

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 50



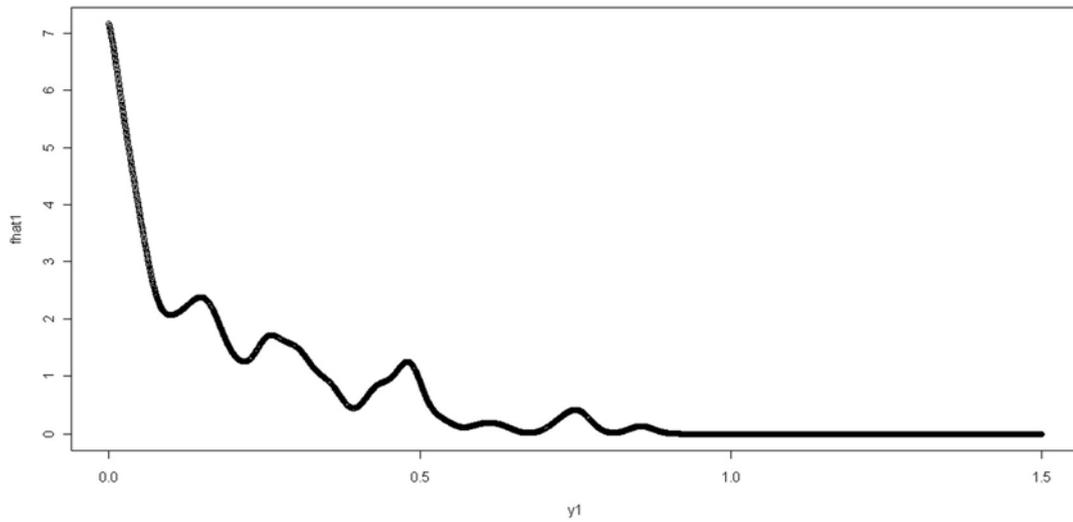
ภาพที่ 4.10

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 100



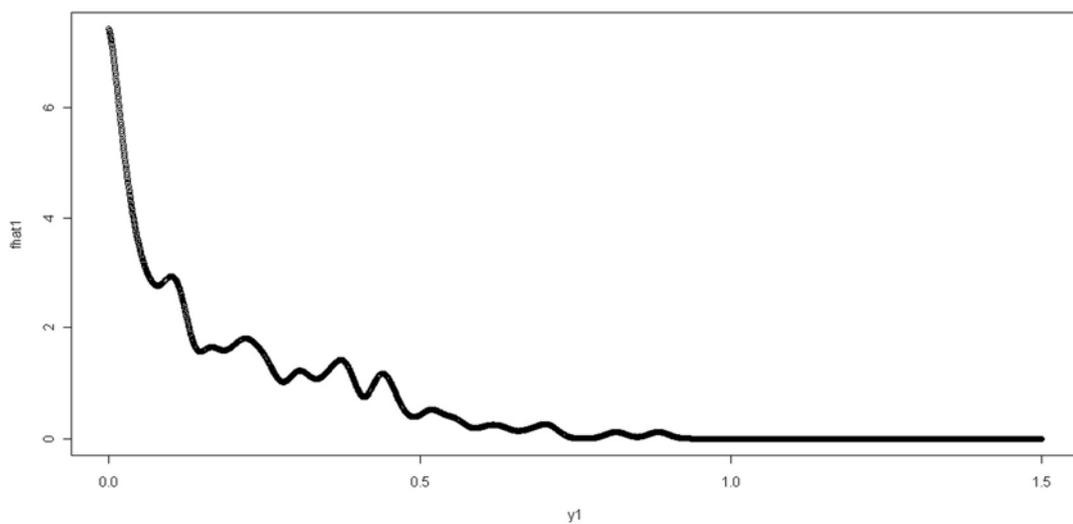
ภาพที่ 4.11

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 150



ภาพที่ 4.12

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 200



จากภาพที่ 4.3 ถึงภาพที่ 4.12 กราฟความหนาแน่นที่ประมาณขึ้นโดยใช้ค่าแบนวิตซ์ที่ได้จากการเลือกแบนวิตซ์ด้วยวิธี Rules of thumb จะได้กราฟความหนาแน่นที่มีความเรียบอย่างเหมาะสมและสามารถสะท้อนให้เห็นถึงลักษณะที่แท้จริงของข้อมูลได้เป็นอย่างดี กราฟความหนาแน่นที่ประมาณขึ้นโดยใช้ค่าแบนวิตซ์ที่ได้จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation จะได้กราฟความหนาแน่นที่มีความเรียบน้อยกว่าปกติ เกิดการขึ้นลงของกราฟบ่อยครั้ง เกิดรอยหยักบนกราฟเป็นจำนวนมากและมีความเรียบน้อยที่สุดในทุกขนาดตัวอย่าง (กราฟความหนาแน่นของฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 ที่ประมาณขึ้นโดยใช้ค่าแบนวิตซ์ที่ได้จากการเลือกแบนวิตซ์ทุกวิธี จำแนกตามขนาดตัวอย่าง แสดงในภาคผนวก ง)

4.1.2 กรณีศึกษาที่ 2

ผลการเปรียบเทียบประสิทธิภาพของวิธีการเลือกแบนวิดท์ แสดงได้ดังตารางที่ 2

ตารางที่ 2

ผลการเปรียบเทียบประสิทธิภาพของวิธีการเลือกแบนวิดท์ 6 วิธี กรณีตัวแปรสุ่ม
เกิดจากฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2

ขนาดตัวอย่าง (n)	วิธีการเลือกแบนวิดท์	MISE	แบนวิดท์เฉลี่ย
30	DPI	0.08591976	0.1037488
	ROT	0.06072683	0.1465575
	LSCV	0.1109543**	0.1098405
	SROT	0.08023053	0.1066623
	BCV	0.05321471*	0.1776107
	STE	0.09605287	0.09336609
50	DPI	0.05611869	0.09240103
	ROT	0.04469938	0.1324114
	LSCV	0.07496323**	0.087177
	SROT	0.05134703	0.1002146
	BCV	0.04336263*	0.1461760
	STE	0.06089471	0.08388595
100	DPI	0.03155954	0.07789808
	ROT	0.02936016	0.1169698
	LSCV	0.04460279**	0.06435416
	SROT	0.03126169	0.08846975
	BCV	0.02744072*	0.1076594
	STE	0.03303121	0.07088106

* ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยน้อยที่สุด

** ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมากที่สุด

ตารางที่ 2 (ต่อ)

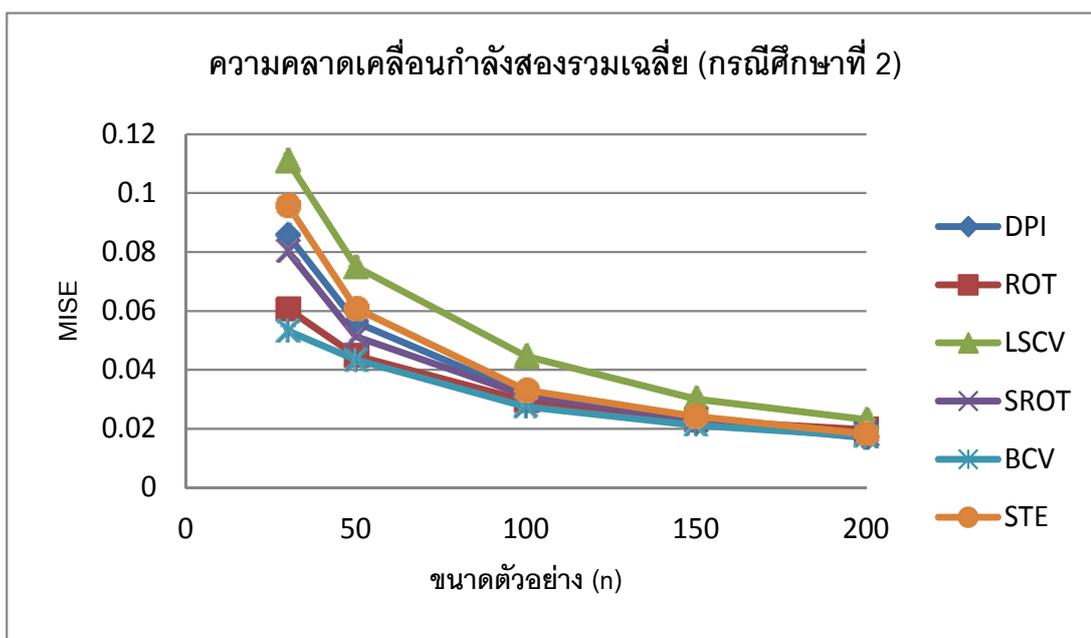
ขนาดตัวอย่าง (n)	วิธีการเลือกแบนด์วิดท์	MISE	แบนด์วิดท์เฉลี่ย
150	DPI	0.02230335	0.06961339
	ROT	0.02285116	0.1079188
	LSCV	0.03002787**	0.05603804
	SROT	0.02304865	0.08284807
	BCV	0.02113676*	0.08950736
	STE	0.02409542	0.06336085
200	DPI	0.01699395*	0.06402614
	ROT	0.01971767	0.1020232
	LSCV	0.02306357**	0.05009436
	SROT	0.01806574	0.07828397
	BCV	0.01751136	0.0805864
	STE	0.01821245	0.05879754

* ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยน้อยที่สุด

** ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมากที่สุด

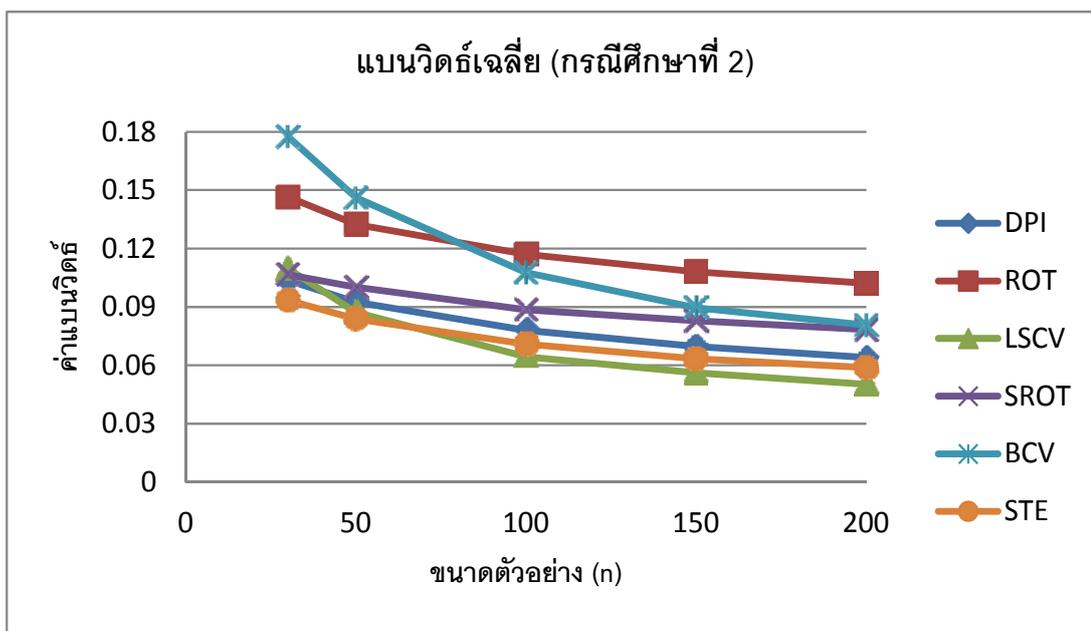
ภาพที่ 4.13

การเปรียบเทียบค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย กรณีตัวแปรสุ่ม
เกิดจากฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2



จากตารางที่ 2 และภาพที่ 4.13 การเลือกแบนด์วิธด้วยวิธี Biased cross-validation จะให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นน้อยที่สุดในเกือบทุกขนาดตัวอย่าง การเลือกแบนด์วิธด้วยวิธี Silverman's rules of thumb และ วิธี Rules of thumb ให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นใกล้เคียงกันในทุกขนาดตัวอย่างและให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นที่ดีรองจากวิธี Biased cross-validation ส่วนวิธี Least squares cross-validation ให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นสูงที่สุดในทุกขนาดตัวอย่าง และค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

ภาพที่ 4.14
การเปรียบเทียบค่าแบนวิตซ์เฉลี่ย กรณีตัวแปรสุ่มเกิดจาก
ฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2



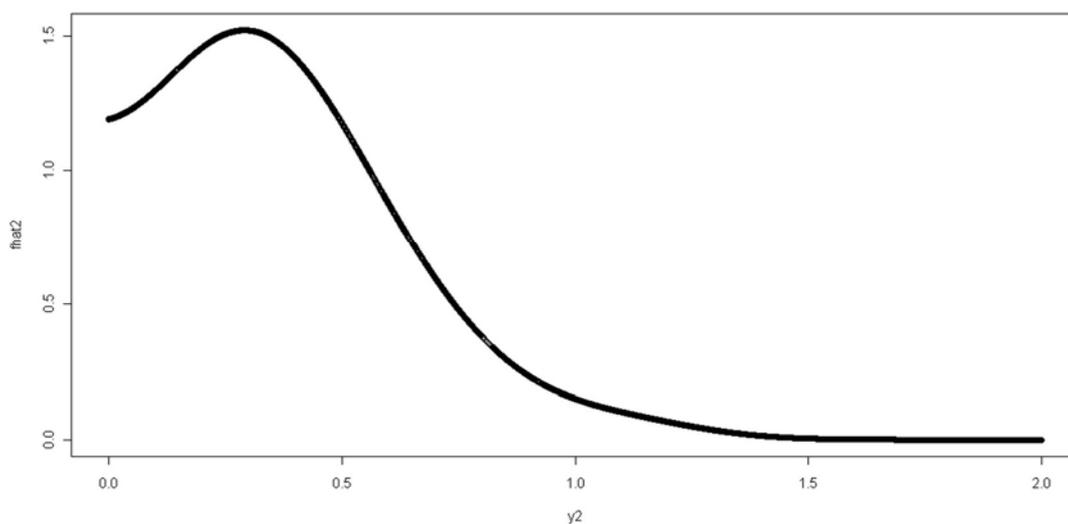
จากตารางที่ 2 และภาพที่ 4.14 การเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation จะให้ค่าแบนวิตซ์เฉลี่ยที่น้อยที่สุดเมื่อขนาดตัวอย่างเท่ากับ 100 , 150 และ 200 เมื่อขนาดตัวอย่างคือ 30 และ 50 วิธี Solve the equation plug-in จะให้ค่าแบนวิตซ์เฉลี่ยที่น้อยที่สุด และเมื่อขนาดตัวอย่างเท่ากับ 30 และ 50 วิธี Biased cross-validation ให้ค่าแบนวิตซ์เฉลี่ยมากที่สุด เมื่อขนาดตัวอย่างเท่ากับ 100 , 150 และ 200 วิธี Rules of thumb ให้ค่าแบนวิตซ์เฉลี่ยมากที่สุด และค่าแบนวิตซ์เฉลี่ยมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

ตัวอย่างกราฟความหนาแน่นที่ได้จากการประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 $f(x) = \frac{5}{4}(1+15x)e^{-5x}$ โดยวิธีการเลือกแบนวิตซ์ที่ทำให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมีค่าน้อยที่สุดเมื่อขนาดตัวอย่างเท่ากับ 30 , 50 , 100 , 150 คือวิธี Biased cross-validation เมื่อขนาดตัวอย่างเท่ากับ 200 คือวิธี Direct plug-in และค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมากที่สุดในทุกขนาดตัวอย่างคือวิธี Least squares cross-validation แสดงได้ดังต่อไปนี้

กราฟความหนาแน่นที่ได้จากการประมาณความหนาแน่นโดยใช้ค่าแบนวิดท์ที่ได้จากวิธี Biased cross-validation เมื่อขนาดตัวอย่างคือ 30 , 50 , 100 , 150 แสดงได้ดังภาพที่ 4.15 ถึงภาพที่ 4.18 ตามลำดับ และวิธี Direct plug-in เมื่อขนาดตัวอย่างคือ 200 แสดงได้ดังภาพที่ 4.19

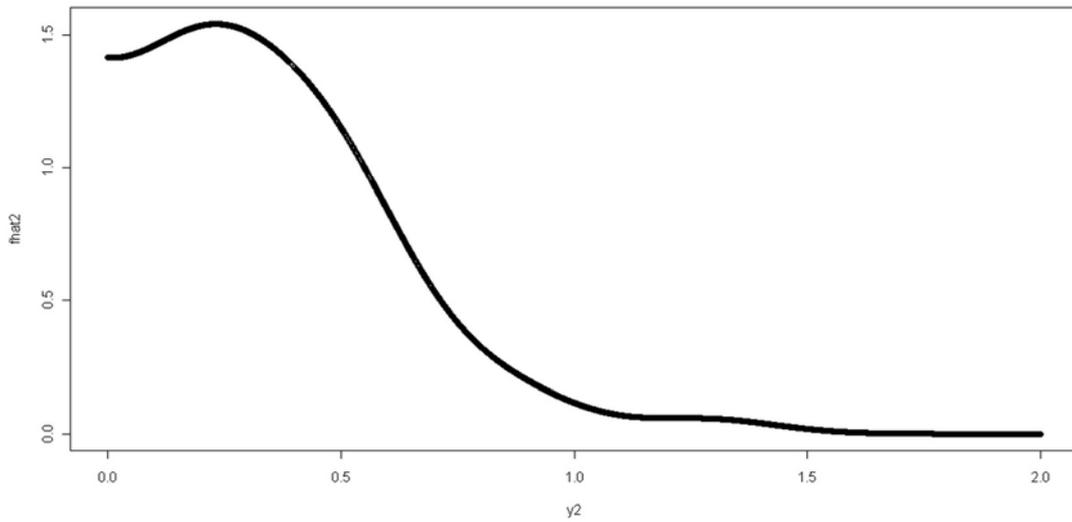
ภาพที่ 4.15

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิดท์จากการเลือกแบนวิดท์ด้วยวิธี Biased cross-validation เมื่อขนาดตัวอย่างคือ 30



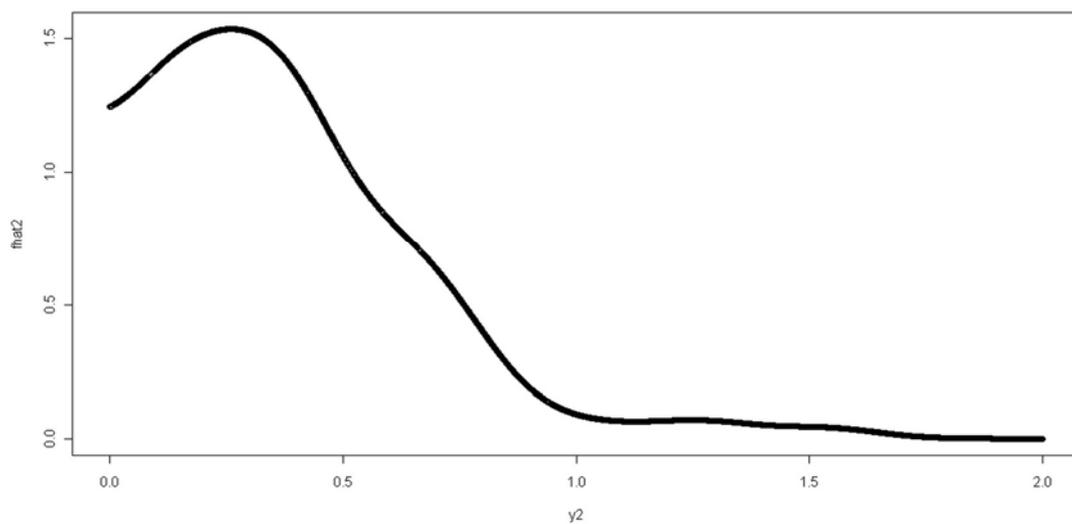
ภาพที่ 4.16

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือก
แบนวิตซ์ด้วยวิธี Biased cross-validation เมื่อขนาดตัวอย่างคือ 50



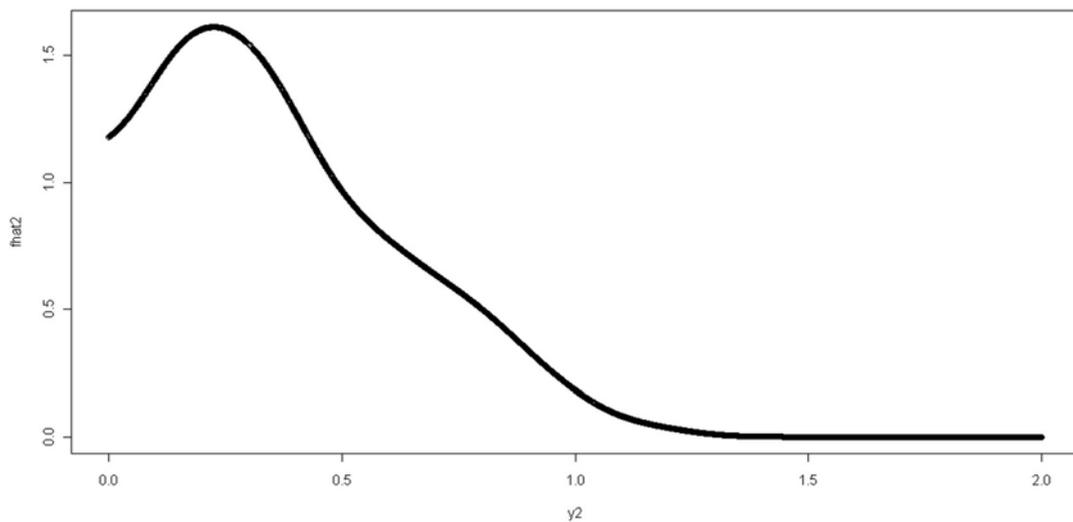
ภาพที่ 4.17

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือก
แบนวิตซ์ด้วยวิธี Biased cross-validation เมื่อขนาดตัวอย่างคือ 100



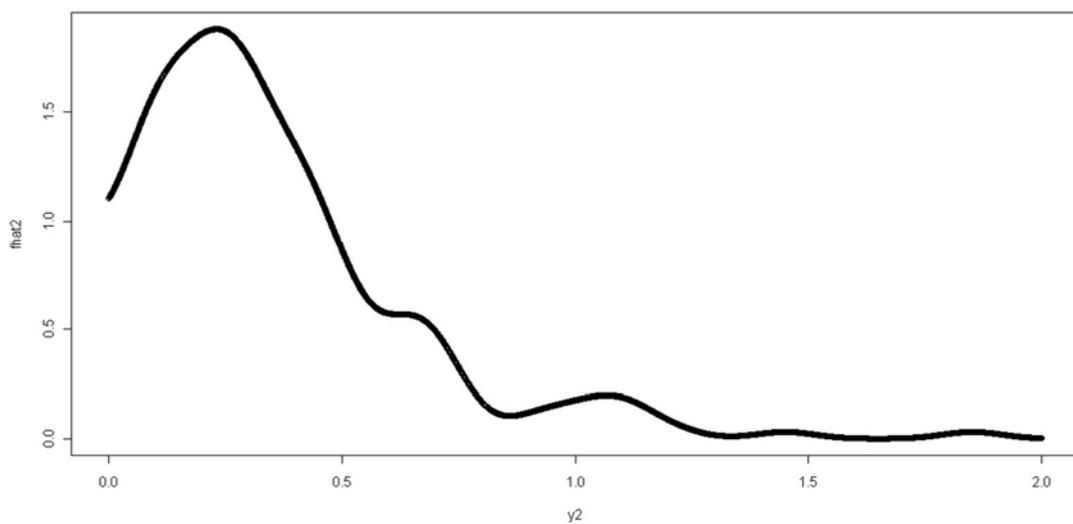
ภาพที่ 4.18

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Biased cross-validation เมื่อขนาดตัวอย่างคือ 150



ภาพที่ 4.19

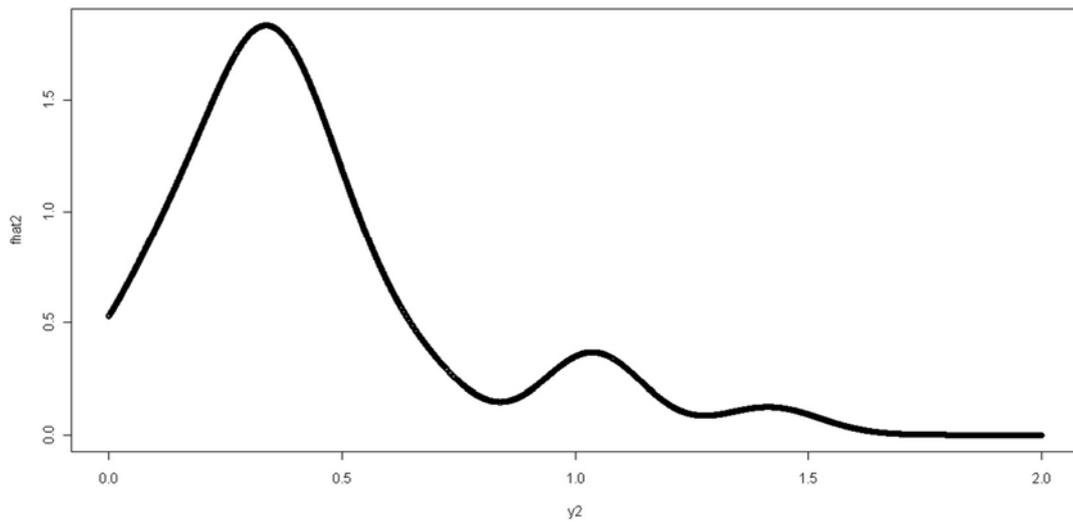
การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Direct plug-in เมื่อขนาดตัวอย่างคือ 200



กราฟความหนาแน่นที่ได้จากการประมาณความหนาแน่นโดยใช้ค่าแบนวิตซ์ที่ได้จากวิธี Least squares cross-validation จำแนกตามขนาดตัวอย่าง แสดงได้ดังภาพที่ 4.20 ถึง ภาพที่ 4.24

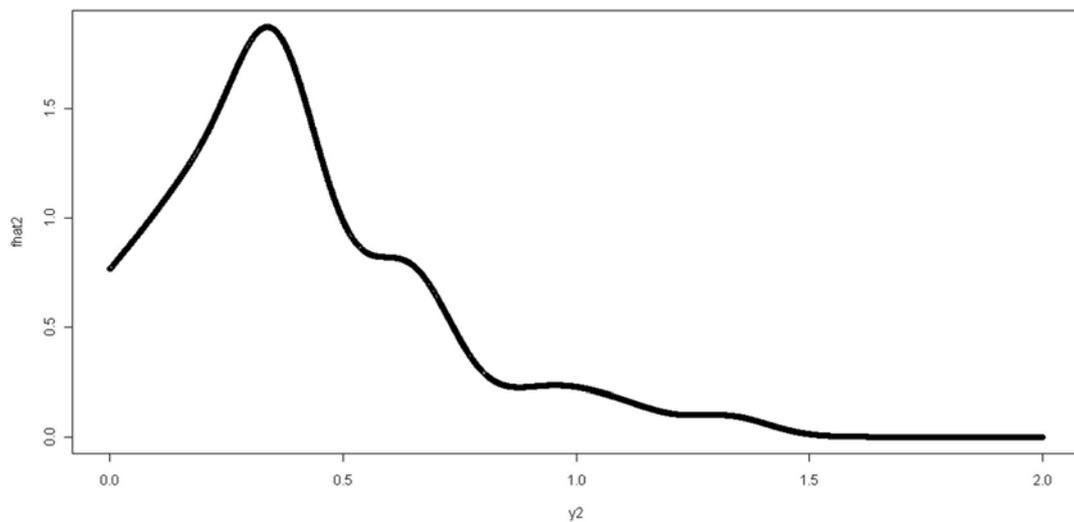
ภาพที่ 4.20

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 30



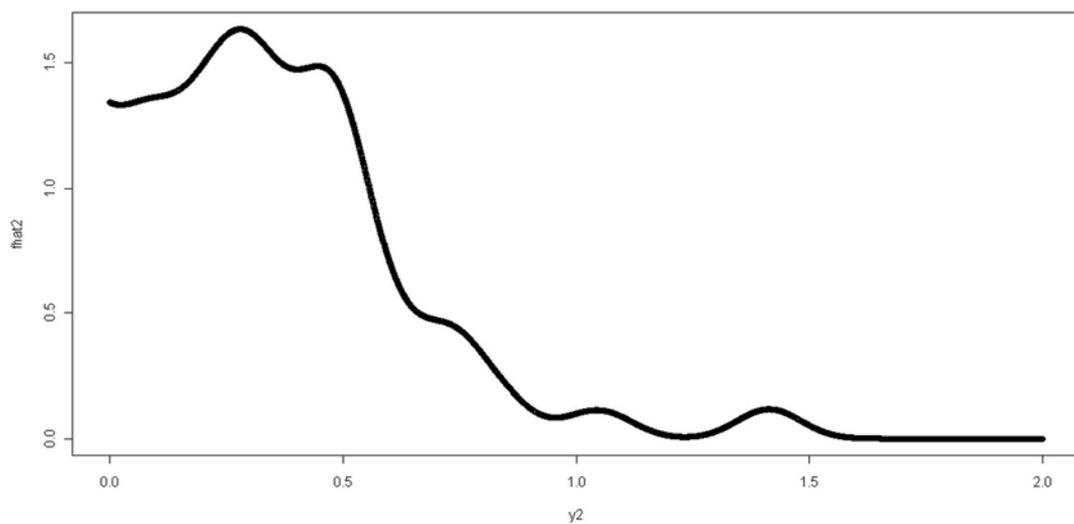
ภาพที่ 4.21

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 50



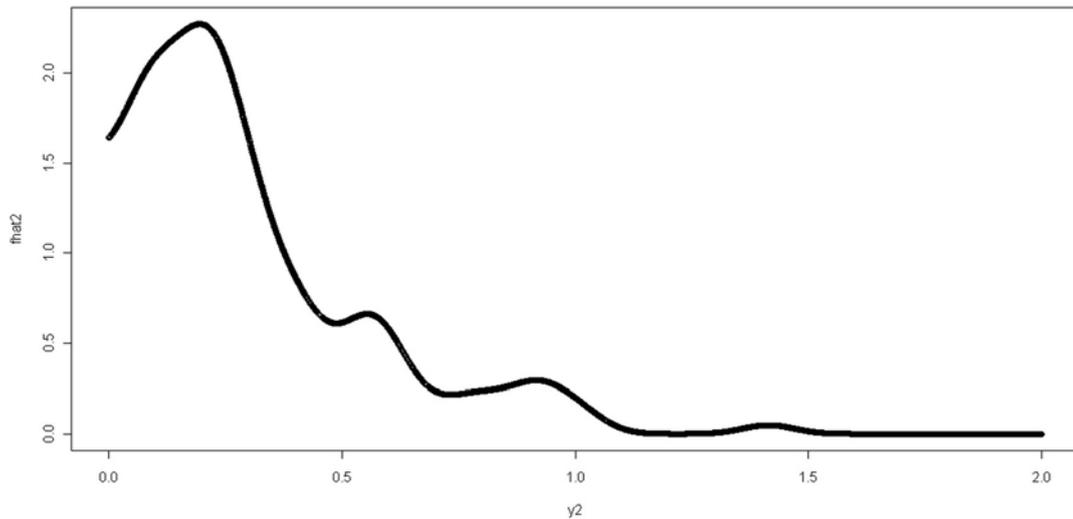
ภาพที่ 4.22

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 100



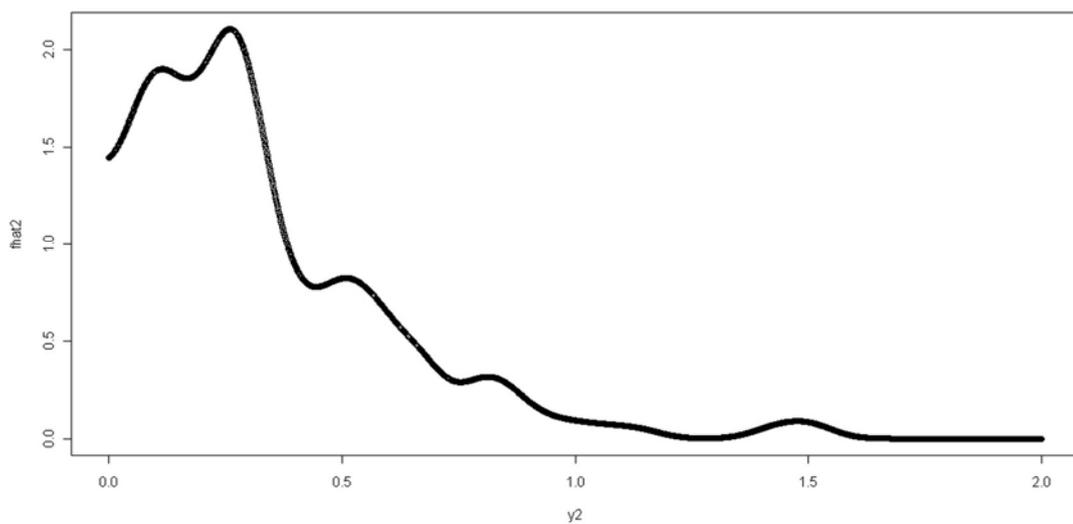
ภาพที่ 4.23

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 150



ภาพที่ 4.24

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี Least squares cross-validation เมื่อขนาดตัวอย่างคือ 200



จากภาพที่ 4.15 ถึงภาพที่ 4.24 กราฟความหนาแน่นที่ประมาณขึ้นโดยใช้ค่าแบน-
วิดธ์ที่ได้จากการเลือกแบนวิดธ์ด้วยวิธี Least squares cross-validation จะได้กราฟความ-
หนาแน่นที่มีความเรียบน้อยที่สุด นั่นคือ เกิดการขึ้นลงของกราฟบ่อยครั้งในทุกขนาดตัวอย่าง
กราฟความหนาแน่นที่ประมาณขึ้นโดยใช้ค่าแบนวิดธ์ที่ได้จากการเลือกแบนวิดธ์ด้วยวิธี Biased
cross-validation จะได้กราฟความหนาแน่นที่มีความเรียบอย่างเหมาะสม สามารถสะท้อนให้เห็น
ถึงลักษณะที่แท้จริงของข้อมูลได้เป็นอย่างดี (กราฟความหนาแน่นของฟังก์ชันความหนาแน่นของ-
ความน่าจะเป็นรูปแบบที่ 2 ที่ประมาณขึ้นโดยใช้ค่าแบนวิดธ์ที่ได้จากการเลือกแบนวิดธ์ทุกวิธี
จำแนกตามขนาดตัวอย่าง แสดงในภาคผนวก ง)

4.2 ผลการวิจัยของการศึกษาข้อมูลเชิงตัวเลข

ข้อมูลจริงที่ใช้ในการวิจัย คือข้อมูลระยะเวลา (วัน) ที่ใช้ในการบำบัดรักษาผู้ป่วย-ควบคุมที่มีอาการทางจิตจำนวน 86 คน แสดงได้ดังตารางที่ 3 ข้อมูลชุดดังกล่าวเป็นข้อมูลของ Copas & Fryer (1980) ถูกใช้แสดงตัวอย่างการประมาณความหนาแน่นด้วยวิธีเคอร์เนล กราฟความหนาแน่นแสดงได้ดังภาพที่ 2.9(a) และภาพที่ 2.9(b) (Silverman, 1986, pp. 18) ภาพที่ 2.9 (a) เป็นการประมาณความหนาแน่นโดยใช้ค่าแบนวิดธ์เท่ากับ 20 และภาพที่ 2.9 (b) เป็นการประมาณความหนาแน่นโดยใช้ค่าแบนวิดธ์คือ 60 ซึ่งผู้วิจัยได้นำกราฟความหนาแน่นทั้งสองมาแสดงได้ดังภาพที่ 4.25

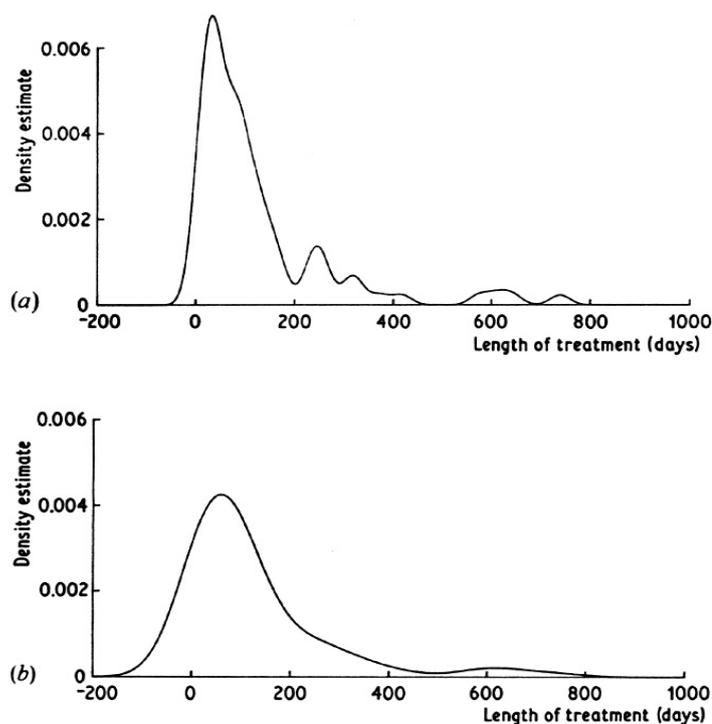
ตารางที่ 3

ข้อมูลระยะเวลา (วัน) ที่ใช้ในการบำบัดรักษาผู้ป่วยควบคุมที่มีอาการทางจิตจำนวน 86 คน

1	25	40	83	123	256
1	27	49	84	126	257
1	27	49	84	129	311
5	30	54	84	134	314
7	30	56	90	144	322
8	31	56	91	147	369
8	31	62	92	153	415
13	32	63	93	163	573
14	34	65	93	167	609
14	35	65	103	175	640
17	36	67	103	228	737
18	37	75	111	231	
21	38	76	112	235	
21	39	79	119	242	
22	39	82	122	256	

ภาพที่ 4.25

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นสำหรับข้อมูลจริงด้วยวิธีเคอร์เนล



จากภาพที่ 4.25 พบว่า การประมาณความหนาแน่นด้วยวิธีเคอร์เนลไม่สามารถประมาณความหนาแน่นบริเวณขอบเขตได้อย่างถูกต้อง เนื่องจากลักษณะกราฟความหนาแน่นที่ได้นั้นเลยออกไปในบริเวณที่ไม่มีข้อมูลตกอยู่ ผู้วิจัยจึงนำข้อมูลชุดดังกล่าวมาประยุกต์ใช้กับงานวิจัยโดยประมาณความหนาแน่นด้วยตัวประมาณความหนาแน่นของ Zhang และคณะ (1999) เพื่อทำการแก้ไขปัญหาการประมาณความหนาแน่นบริเวณขอบเขต

ลักษณะกราฟความหนาแน่นที่ได้จากการประมาณความหนาแน่นด้วยวิธีของ Zhang และคณะ (1999) โดยใช้ค่าแบนวิดธ์ที่ได้จากวิธีการเลือกแบนวิดธ์ 6 วิธีคือ วิธี Direct plug-in (DPI) วิธี Rules of thumb (ROT) วิธี Least squares cross-validation (LSCV) วิธี Silverman's rules of thumb (SROT) วิธี Biased cross-validation (BCV) และ วิธี Solve the equation plug-in (STE) แสดงได้ดังภาพที่ 4.26 ถึง ภาพที่ 4.31 ตามลำดับ และค่าแบนวิดธ์ของแต่ละวิธีการเลือกแบนวิดธ์ที่คำนวณได้จากข้อมูลจริงชุดดังกล่าว แสดงค่าในตารางที่ 4

ตารางที่ 4

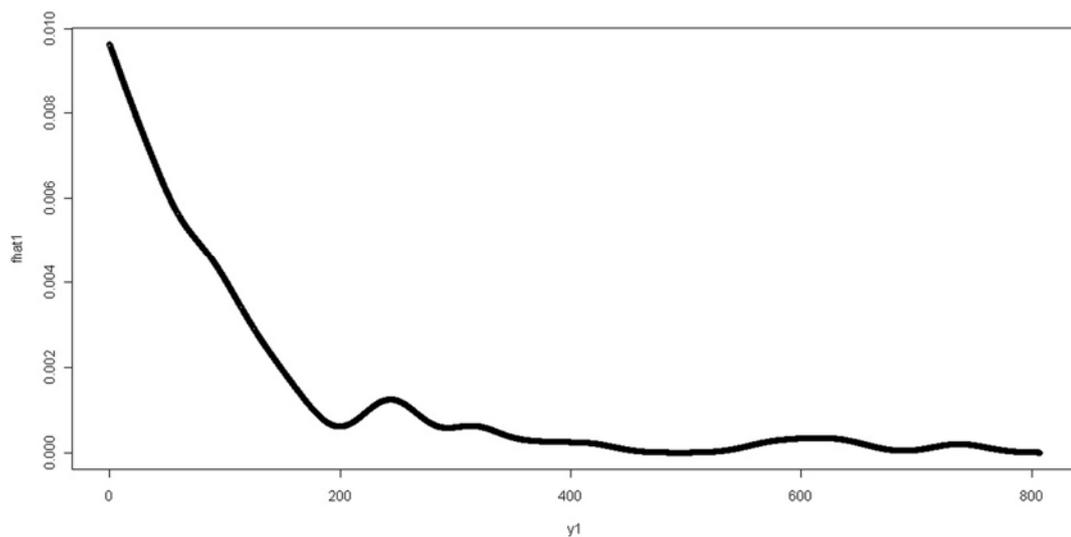
ค่าแบนด์วิดท์ที่ใช้ในการประมาณความหนาแน่นสำหรับข้อมูลจริง

วิธีการเลือกแบนด์วิดท์	ค่าแบนด์วิดท์
DPI	23.15698
ROT	63.8208
LSCV	15.90555
SROT	30.38202
BCV	33.42486
STE	19.40324

ภาพที่ 4.26

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นสำหรับข้อมูลจริง
ด้วยวิธีของ Zhang และคณะ ใช้ค่าแบนด์วิดท์จากการเลือกแบนด์วิดท์ด้วยวิธี

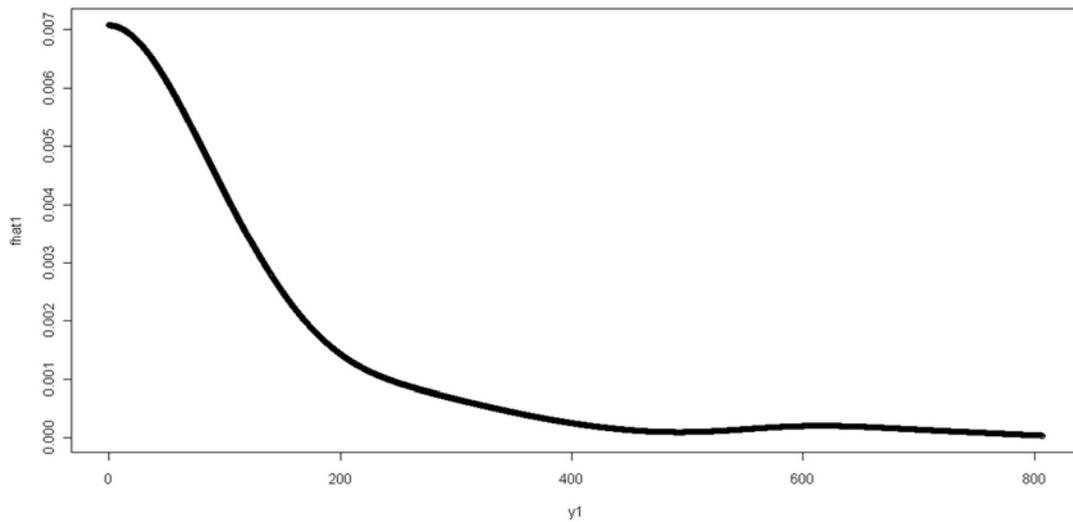
Direct plug-in



ภาพที่ 4.27

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นสำหรับข้อมูลจริง
ด้วยวิธีของ Zhang และคณะ ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี

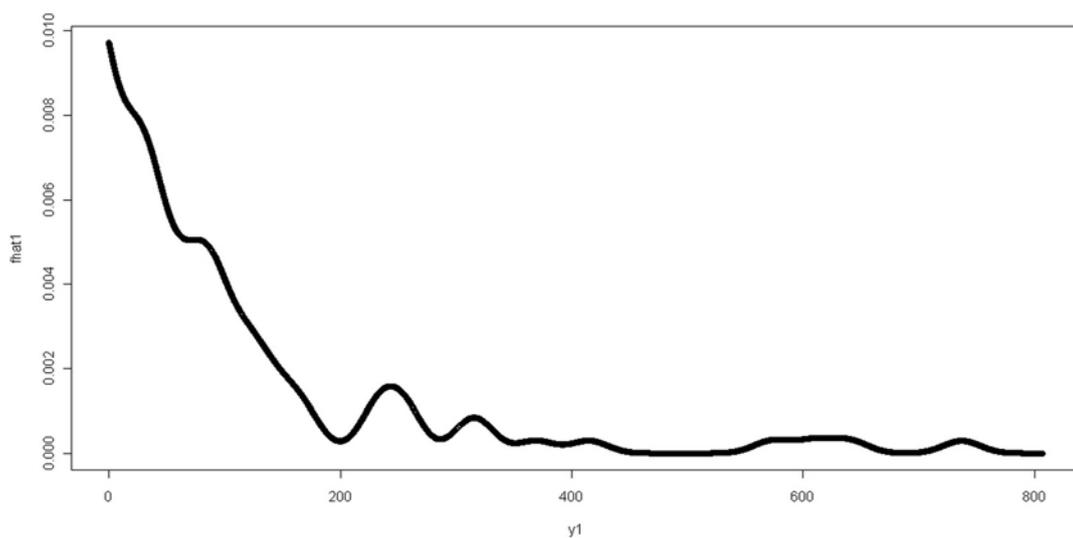
Rules of thumb



ภาพที่ 4.28

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นสำหรับข้อมูลจริง
ด้วยวิธีของ Zhang และคณะ ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี

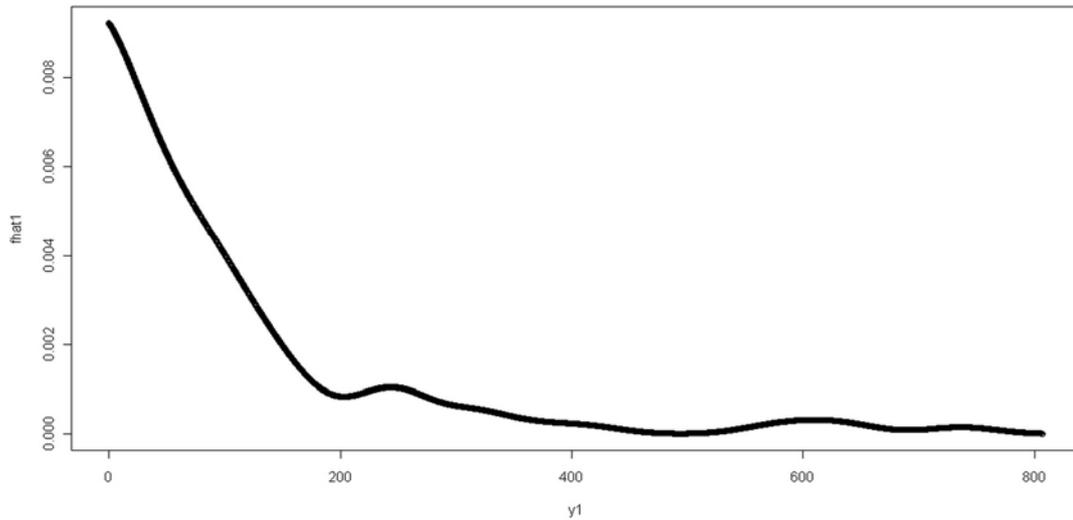
Least squares cross-validation



ภาพที่ 4.29

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นสำหรับข้อมูลจริง
ด้วยวิธีของ Zhang และคณะ ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี

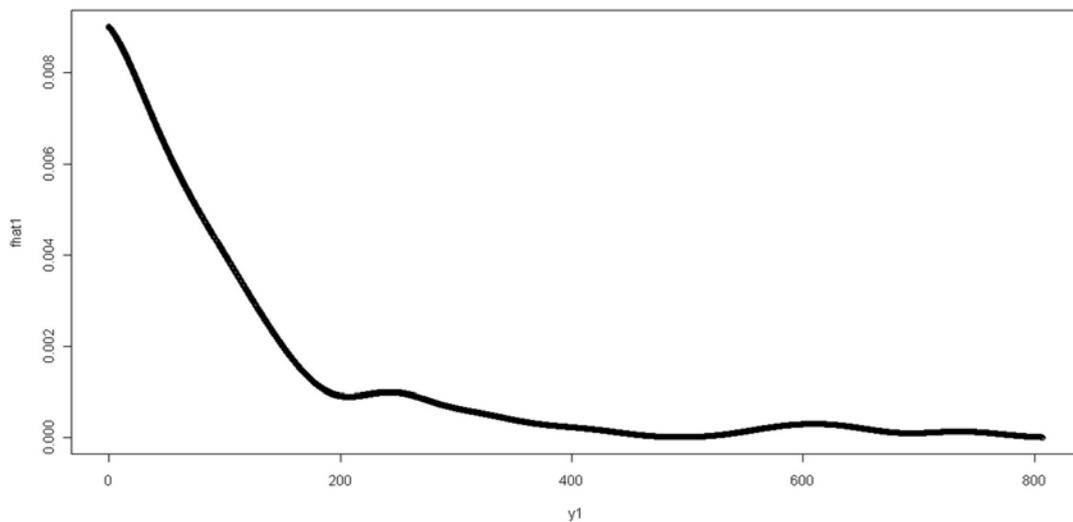
Silverman's rules of thumb



ภาพที่ 4.30

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นสำหรับข้อมูลจริง
ด้วยวิธีของ Zhang และคณะ ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี

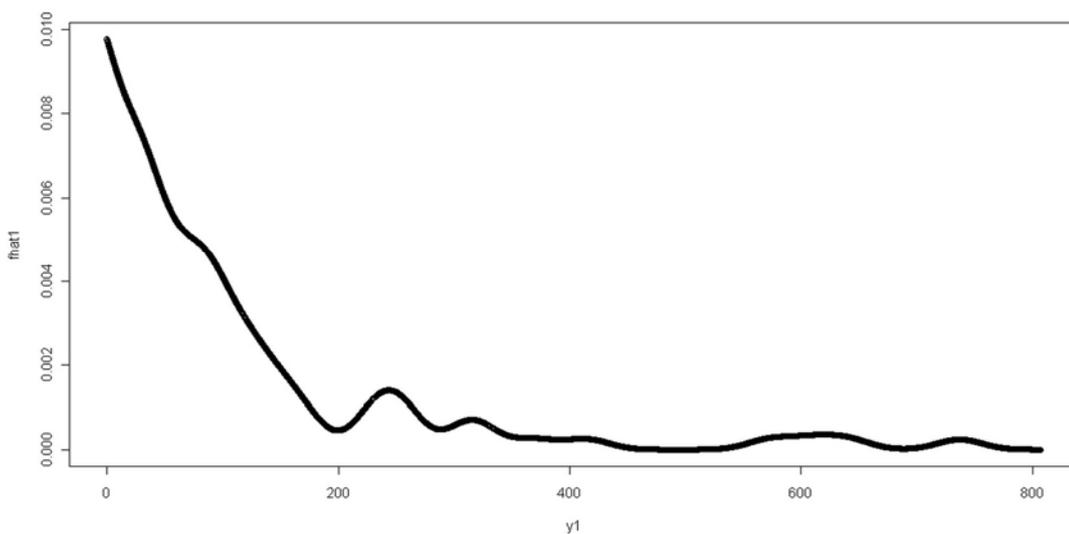
Biased cross-validation



ภาพที่ 4.31

การประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นสำหรับข้อมูลจริง
ด้วยวิธีของ Zhang และคณะ ใช้ค่าแบนวิตซ์จากการเลือกแบนวิตซ์ด้วยวิธี

Solve the equation plug-in



จากตารางที่ 4 และภาพที่ 4.26 ถึงภาพที่ 4.31 วิธีการเลือกแบนวิตซ์ที่ให้ค่าแบนวิตซ์มากที่สุดและมากกว่าวิธีอื่น ๆ อย่างชัดเจน คือวิธี Rules of thumb จึงส่งผลให้ลักษณะกราฟความหนาแน่นมีความเรียบมากกว่าวิธีอื่น ๆ ส่วนวิธี Least squares cross-validation ให้ค่าแบนวิตซ์น้อยที่สุด ทำให้กราฟความหนาแน่นเกิดการขึ้นลงบ่อยครั้งและมีความเรียบน้อยกว่าวิธีอื่น ๆ