บทที่ 2

ทฤษฎีที่เกี่ยวข้องในการวิจัย

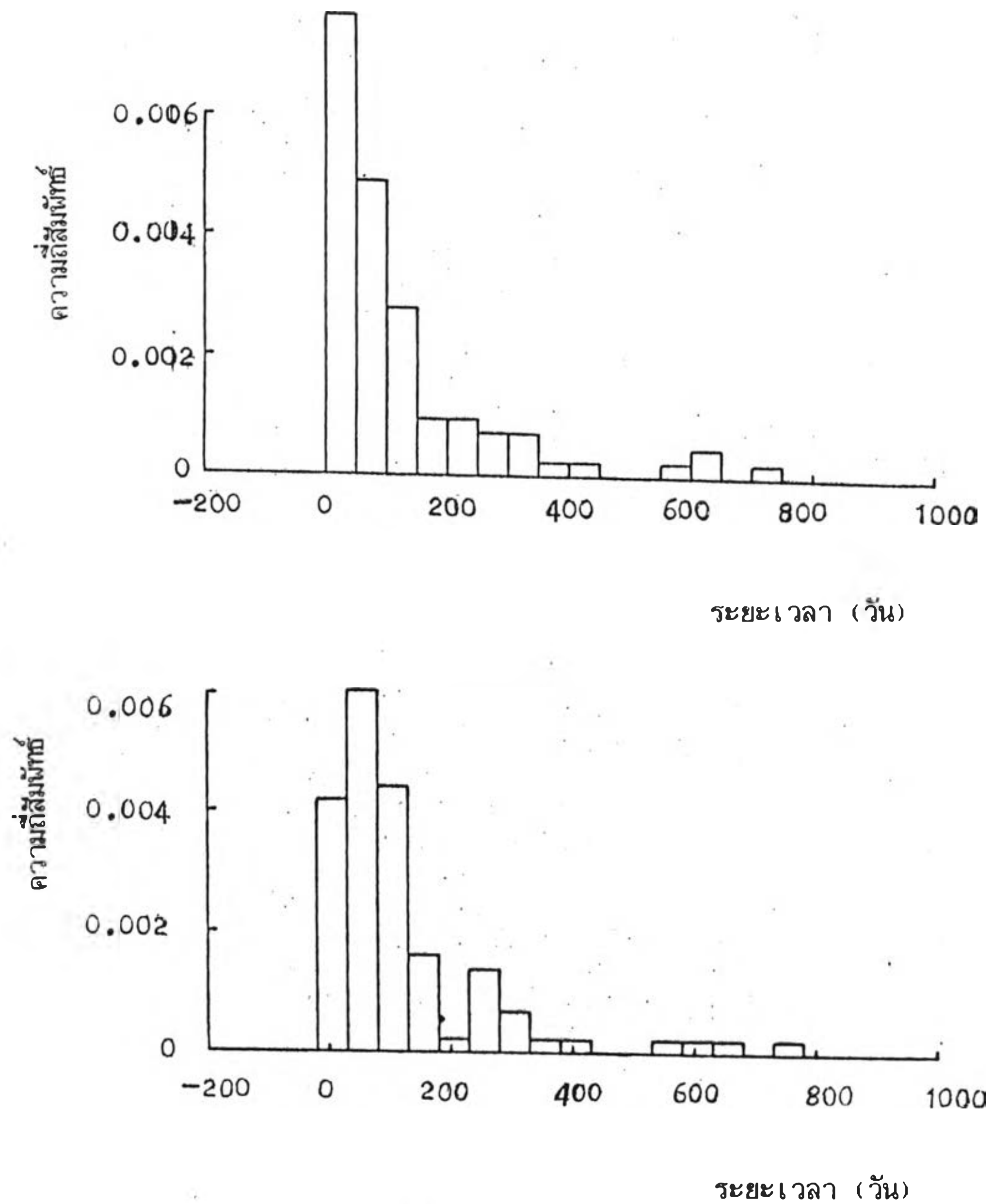
วิธีที่ใช้กันมากในการประมาณการแจกแจงของประชากรซึ่งเป็นวิธีเก่าที่สุดคือวิธีฮิสโตแกรม (Histogram Method)[๕] การสร้างฮิสโตแกรม เริ่มจากการกำหนดจุดกำเนิดและการกำหนดความกว้างของช่วง จุดกำเนิดมักกำหนดจากค่าต่ำสุดของข้อมูลที่รวบรวมได้ ส่วนความกว้างของช่วง ประมาณจากผลหารระหว่างผลต่างของค่าสูงสุดกับค่าต่ำสุดของข้อมูล และจำนวนช่วง ความกว้างของช่วงที่เหมาะสมจะทำให้ได้รูปแบบการแจกแจงที่ใกล้เคียงของประชากร ถ้าใช้จำนวนช่วงน้อยเกินไป (คือแต่ละช่วงกว้างเกินไป) จะทำให้ไม่เห็นรูปแบบการแจกแจงที่ถูกต้อง แต่ถ้าจำนวนช่วงมากเกินไป (คือแต่ละช่วงแคบเกินไป) ก็จะทำให้เกิดความสับสน เพราะช่วงหนึ่ง ๆ อาจมีค่าสังเกตเพียงค่าเดียวหรือไม่มีเลยก็ได้ ดังนั้น จึงมีข้อเสนอแนะให้ใช้จำนวนช่วงระหว่าง 8 - 20 ช่วง หรือประมาณจำนวนช่วงจาก $\sqrt{N}$ เมื่อ N เป็นขนาดของกลุ่มตัวอย่าง[๖]

นอกจากความกว้างของช่วงแล้ว จุดกำเนิด (หรือจุดเริ่มต้น) ก็มีผลมากจากการสร้างฮิสโตแกรมโดยใช้ข้อมูลชุดเดียวกัน และใช้ความกว้างของแต่ละช่วงเท่ากัน แต่จุดที่เริ่มต้นของฮิสโตแกรมต่างกันจะได้ผลแตกต่างกันดังรูปที่ 2.1

---

[๕]B.W. Silverman, Density Estimation for Statistics and Data Analysis (London : Chapman and Hall, 1986), p. 7.

[๖]Wilfrid J. Dixon, Introduction to Statistical Analysis (McGraw - Hill, 1983), p. 8.



รูปที่ 2.1 ฮิสโตแกรมแสดงระยะเวลา (วัน) ในการทดลองกับคนไข้กลุ่มควบคุมที่ใช้ศึกษาเรื่องการฆ่าตัวตาย[7]

---

[7]B.W. Silverman, Density Estimation for Statistics and Data Analysis (London : Chapman and Hill, 1986), p. 11.

จากฮิสโตแกรมในรูปที่ 2.1 เป็นฮิสโตแกรมที่สร้างขึ้นโดยใช้ข้อมูลชุดเดียวกัน ความกว้างของช่วงเท่ากัน แต่จุดเริ่มต้นต่างกัน ซึ่งทำให้ได้ฮิสโตแกรมที่ให้ความหมายแตกต่างกัน

ดังนั้น การประมาณการแจกแจงโดยใช้การสร้างฮิสโตแกรม แม้จะรู้สึกว่าง่าย แต่ปัญหาก็มีเกิดขึ้นดังที่ได้กล่าวมาแล้ว นอกจากการประมาณจากฮิสโตแกรมแล้ว ยังมีวิธีการประมาณฟังก์ชันความหนาแน่นด้วยวิธีอื่น ๆ อีก แต่ทุกวิธีของการประมาณใช้วิธีพื้นฐานที่ต่างกัน 2 อย่าง คือ

1. วิธีของฟังก์ชันถ่วงน้ำหนัก (Method of weight function) วิธีนี้ฟังก์ชันความหนาแน่นที่ประมาณได้สร้างขึ้นโดยใช้ฟังก์ชันถ่วงน้ำหนัก (weight function) ซึ่งต้องสอดคล้องกับเงื่อนไขบางประการ เช่น ต้องเป็นฟังก์ชันสมมาตร เป็นต้น

2. วิธีการกระจายของ f(x) (Method of "expanding of f(x)") วิธีนี้ฟังก์ชันความหนาแน่นที่ประมาณสร้างขึ้นโดยใช้ฟังก์ชันออร์โธนอร์มัล (Orthonormal function)[8]

โดยอาศัยวิธีของฟังก์ชันถ่วงน้ำหนัก วิธีหนึ่งที่นิยมใช้ในการประมาณฟังก์ชันความหนาแน่นกันมาก คือ วิธีเคอร์เนล (Kernel Method) วิธีนี้เป็นวิธีที่ได้รับการสนับสนุนให้เลือกใช้เป็นอันดับแรก โดยเฉพาะเมื่อต้องการประมาณฟังก์ชันความหนาแน่นของหนึ่งตัวแปร เพราะวิธีนี้เป็นวิธีที่สามารถประยุกต์ได้อย่างกว้างขวาง[9] วิธีประมาณฟังก์ชันความหนาแน่นโดยวิธีเคอร์เนล ฟังก์ชันถ่วงน้ำหนัก คือ ฟังก์ชันเคอร์เนล โดยทั่วไป (ไม่เสมอไป) ฟังก์ชันเคอร์เนลจะเป็นฟังก์ชันสมมาตร และสอดคล้องกับเงื่อนไขของการเป็นฟังก์ชันความน่าจะเป็นหนาแน่น กล่าวคือ ถ้า K เป็นฟังก์ชันเคอร์เนล ฟังก์ชัน K

---

[8]M.A. Mirzahmedov, <u>On Some Properties of Density Estimation</u> (Hungary : Budapest, 1974), p. 535.

[9]B.W. Silverman, <u>Density Estimation for Statistics and Data Analysis</u> (London : Chapman and Hill, 1986), p. 34.

ต้องสอดคล้องกับเงื่อนไขที่ว่า[10]

$$\int_{-\infty}^{\infty} K(x)\, dx = 1$$

การประมาณฟังก์ชันความหนาแน่นของประชากรโดยวิธีเคอร์เนล ประมาณได้จากสูตร[11]

$$\hat{f}(x) = (1/nh)\sum_{i=1}^{n} K[(x-X_i) / h]$$

โดยที่ $\hat{f}$ เป็นฟังก์ชันความหนาแน่นที่ประมาณได้

n เป็นขนาดของกลุ่มตัวอย่าง

h เป็นความกว้าง

K เป็นฟังก์ชันเคอร์เนล

X_i เป็นค่าสังเกต (ค่าของข้อมูลตัวอย่าง)

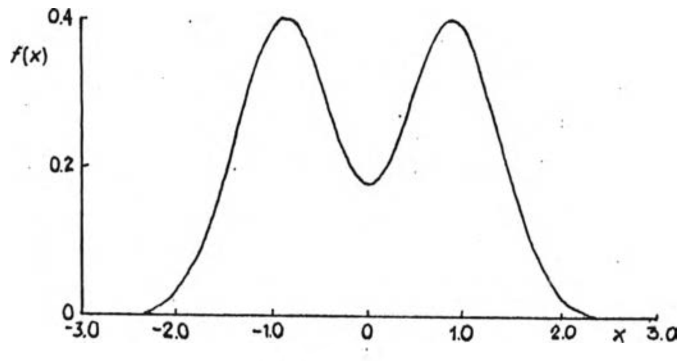
การประมาณฟังก์ชันความหนาแน่นของประชากรดังวิธีที่กล่าวข้างต้น ฟังก์ชันที่ประมาณได้คือ ฟังก์ชัน $\hat{f}$ จากสูตรจะเห็นได้ว่า ค่าของ $\hat{f}$ ขึ้นกับค่าของ h ฟังก์ชัน K และค่า X_i แต่ตามวิธีของเคอร์เนล ได้สรุปว่า พารามิเตอร์ที่มีผลมากต่อฟังก์ชัน $\hat{f}$ ที่ประมาณได้คือ ค่าของ h โดยที่ h เป็นความกว้างของแต่ละช่วง และจะเรียก h ว่า ความกว้าง (Window width) หรือ พารามิเตอร์ที่ใช้ปรับกราฟ (Smoothing parameter) Silverman (1986) ได้แสดงความเห็นว่า การกำหนดความกว้างของช่วงที่เหมาะสมจะทำให้ได้ฟังก์ชันความหนาแน่นโดยประมาณใกล้เคียงกับฟังก์ชันความหนาแน่นที่แท้จริง[12] ดังรูปที่ 2.2 และ 2.3

---

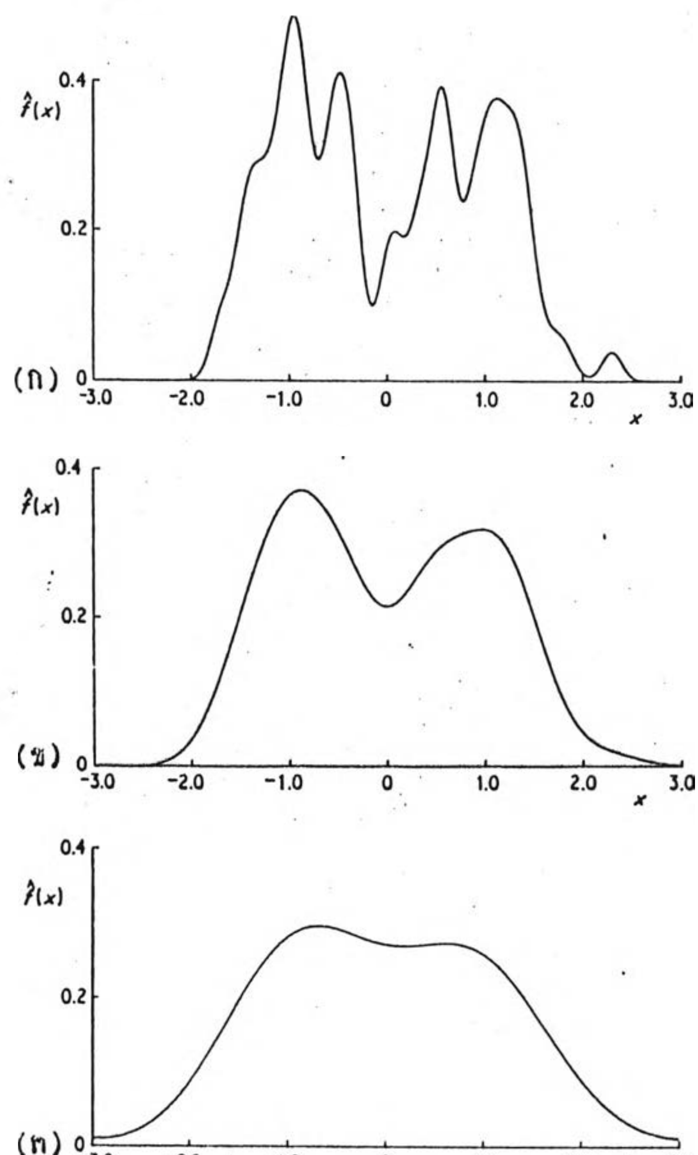
[10]B.W. Silverman, <u>Density Estimation for Statistics and Data Analysis</u> (London : Chapman and Hall, 1986), p. 13.

[11]Ibid., p.15.

[12]Ibid., p.16 - 17.

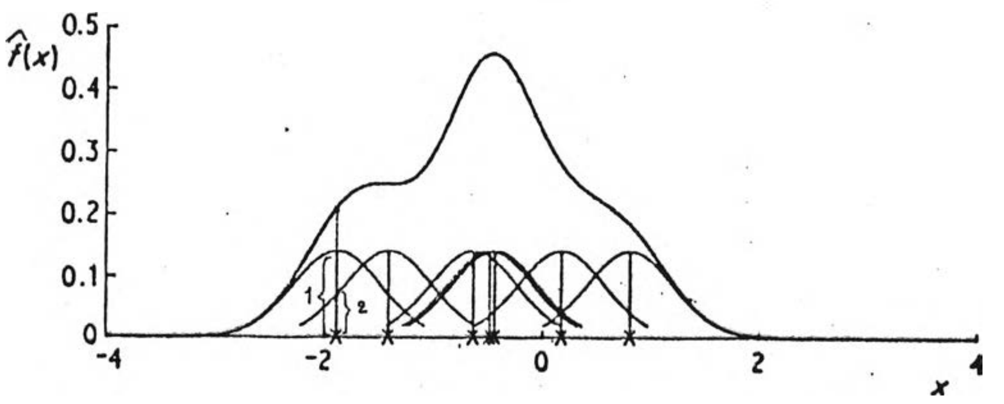


รูปที่ 2.2 ฟังก์ชันความหนาแน่นที่แท้จริง



รูปที่ 2.3 (ก) (ข) และ (ค) เป็นฟังก์ชันความหนาแน่นที่ประมาณจากข้อมูลตัวอย่าง ขนาด 200 โดยวิธีเคอร์เนล ใช้ความกว้าง : (ก) 0.1 (ข) 0.3 (ค) 0.6

รูปที่ 2.2 แสดงฟังก์ชันความหนาแน่นที่แท้จริงของประชากร ส่วนรูปที่ 2.3 (ก) (ข) (ค) เป็นฟังก์ชันความหนาแน่นที่ประมาณจากข้อมูลตัวอย่างขนาด 200 ที่จำลองมาจากประชากรที่แท้จริง คือมีฟังก์ชันความหนาแน่นดังรูปที่ 2.2 ฟังก์ชันเคอร์เนลที่ใช้ในการประมาณคือ $K_2(t) = (1/\sqrt{2\pi})e^{-(1/2)t^2}$ $, -\infty < t < \infty$ แต่ความกว้างที่ใช้แตกต่างกัน รูปที่ 2.3 (ข) ใกล้เคียงกับฟังก์ชันความหนาแน่นที่แท้จริงของประชากรมากกว่า รูปที่ 2.3 (ก) ซึ่งใช้ความกว้างของช่วงเล็กเกินไป หรือรูปที่ 2.3 (ค) ซึ่งใช้ความกว้างของช่วงใหญ่เกินไป จนทำให้ไม่เห็นลักษณะแท้จริงที่มีสองฐานนิยม เพื่อความเข้าใจในเรื่องนี้ดีขึ้น จงพิจารณารูปที่ 2.4 ดังนี้[13]



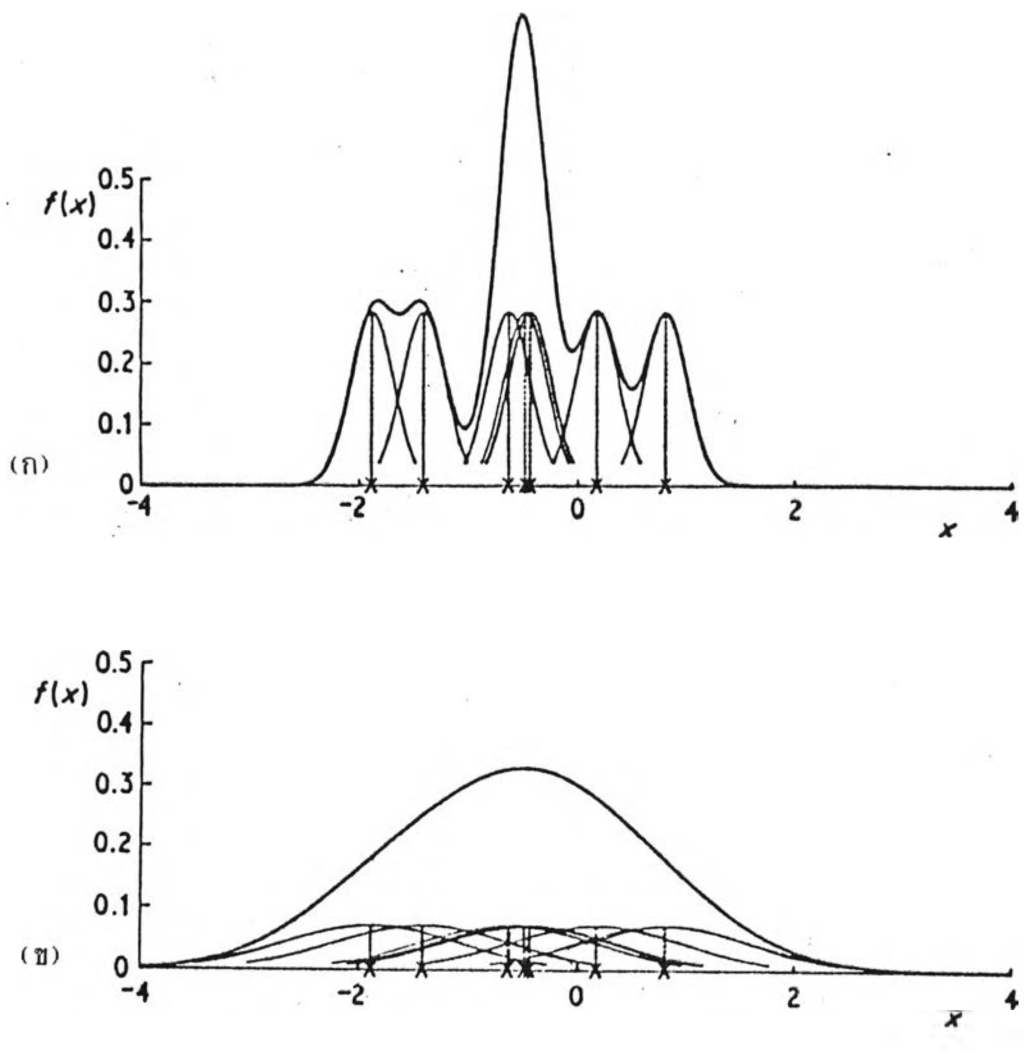
รูปที่ 2.4 การประมาณฟังก์ชันความหนาแน่นโดยวิธีเคอร์เนลโดยแสดงให้เห็นถึงแต่ละเคอร์เนล (ความกว้าง 0.4)

จากรูปที่ 2.4 แสดงการสร้างฟังก์ชัน $\hat{f}$ จากข้อมูลตัวอย่างขนาด 7 ซึ่งขนาดตัวอย่างเป็นขนาดที่เล็ก แต่เพื่อช่วยความเข้าใจได้ง่ายขึ้น จะเห็นได้ว่า ค่าของ $\hat{f}$ ที่ x แต่ละค่าเกิดจากผลรวมของแต่ละเคอร์เนลที่จุดนั้น ๆ เช่นที่ $x = -2$ ค่าของ $\hat{f}$ ที่ $x = -2$ เกิดจากความสูง (1) รวมกับความสูง (2) ฟังก์ชันเคอร์เนลเป็นตัวกำหนดรูปร่างของแต่ละโค้งเล็ก ๆ ที่อยู่ภายใน ขณะที่ความกว้าง h กำหนดความกว้าง ในที่นี้โค้งเล็ก ๆ ต่าง

---

[13] B.W. Silverman, Density Estimation for Statistics and Data Analysis (London : Chapman and Hall, 1986), p. 14.

เป็นโค้งปกติ เพราะใช้ฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนล เพื่อเป็นการแสดงว่าการใช้ฟังก์ชันเคอร์เนลเดียวกัน แต่ใช้ความกว้างต่างกันจะทำให้ได้ฟังก์ชัน $\hat{f}$ ที่ประมาณได้แตกต่างกัน พิจารณารูปที่ 2.5 ดังต่อไปนี้



รูปที่ 2.5 (ก) (ข) การประมาณฟังก์ชันความหนาแน่น โดยวิธีเคอร์เนล โดยแสดงให้เห็นถึงแต่ละเคอร์เนลความกว้าง (ก) 0.2 (ข) 0.8

รูปที่ 2.4 และรูปที่ 2.5 (ก) (ข) ต่างเป็นฟังก์ชันความหนาแน่นที่ประมาณจากข้อมูลตัวอย่างขนาด 7 ใช้ฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนล แต่ใช้ความกว้างของช่วง (h) แตกต่างกัน

ฟังก์ชันเคอร์เนล นอกจากฟังก์ชันความหนาแน่นปกติมาตรฐานตามที่กล่าวมาแล้ว มีอีกหลายแบบ ดังนี้[14]

1) เอพาเนคนิคอฟ (Epanechnikov) ฟังก์ชันเคอร์เนลคือ

$$K(t) = (3/4)[1 - (1/5)t^2] / \sqrt{5}, \quad |t| < \sqrt{5}$$
$$= 0 \quad , \quad t \text{ อื่น ๆ}$$

2) สองฐานนิยม (Biweight) ฟังก์ชันเคอร์เนลคือ

$$K(t) = (5/16)(1 - t^2), \quad |t| < 1$$
$$= 0 \quad , \quad t \text{ อื่น ๆ}$$

3) สามเหลี่ยม (Triangular) ฟังก์ชันเคอร์เนลคือ

$$K(t) = 1 - |t|, \quad |t| < 1$$
$$= 0 \quad , \quad t \text{ อื่น ๆ}$$

แม้ว่าฟังก์ชันเคอร์เนลจะมีหลายแบบ แต่ฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนลที่นิยมใช้กันมาก จากตัวอย่างที่กล่าวมาข้างต้นต่างใช้ฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนลทั้งสิ้น ในปี ค.ศ.1982 โอยาง ซู ปีเตอร์[15] ได้ทำการศึกษาการประมาณฟังก์ชันความหนาแน่นโดยศึกษาวิธีการประมาณ 3 วิธี คือ วิธีฮิสโตแกรม (Histogram Method) วิธีควอนไทล์ (Quantile Method) และวิธีเคอร์เนล ซึ่งเป็นการประมาณฟังก์ชันความหนาแน่น $\hat{f}$ โดยใช้ตัวประมาณเคอร์เนล การศึกษาครั้งนี้ฟังก์ชันเคอร์เนลที่ใช้ในการศึกษาคือ ฟังก์ชันความหนาแน่นปกติมาตรฐาน

---

[14] B.W. Silverman, Density Estimation for Statistics and Data Analysis (London : Chapman and Hall, 1986), p. 43.

[15] Ouyang Soo Peter. "Empirecal Bayes and Density Estimation." PH.D's Thesis, State University of New York At Strong Brook, 1982.

เนื่องจากฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันความหนาแน่นที่นิยมใช้เป็นฟังก์ชันเคอร์เนลกันมาก ในการวิจัยครั้งนี้ ผู้วิจัยจึงเลือกฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนลที่ใช้ในการศึกษาและเลือกฟังก์ชันสามเหลี่ยมมาลองศึกษาเปรียบเทียบด้วย ทั้งนี้เพราะว่าฟังก์ชันสามเหลี่ยมและการแจกแจงของประชากรที่มีการแจกแจงปัวส์ซงพารามิเตอร์เท่ากับ λ กรณี λ มีค่าน้อย ๆ มีลักษณะคล้ายกัน และเมื่อ λ มีค่ามากขึ้น การแจกแจงปัวส์ซงจะใกล้เคียงกับการแจกแจงปกติ

ส่วนความกว้างของช่วง ซึ่งเป็นพารามิเตอร์ที่ใช้ปรับกราฟของฟังก์ชันความหนาแน่นที่ประมาณ มีวิธีการกำหนดหลายวิธีโดยที่วิธีการแต่ละวิธีเป็นเพียงแนวทางที่คาดว่าใช้ได้ แต่ยังไม่ได้เป็นที่ยอมรับกันเป็นสากล วิธีการเลือกใช้ค่าพารามิเตอร์ที่ใช้ปรับกราฟควรพิจารณาหลาย ๆ ค่า โดยอาศัยความคิดเบื้องต้นเกี่ยวกับฟังก์ชันความหนาแน่นที่ศึกษา เช่น ถ้าประชากรมีการแจกแจงปกติมีค่าความแปรปรวน σ^2 และใช้ฟังก์ชันความหนาแน่นปกติมาตรฐานเป็นฟังก์ชันเคอร์เนล จะสามารถประมาณค่าที่เหมาะสมได้จากสูตร[16]

$$h_1 = 1.06\,\sigma n^{-(1/5)}$$

h_1 เป็นความกว้างของช่วง σ ประมาณค่าจากส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง แต่ถ้าการแจกแจงของประชากรที่เราสนใจศึกษามีลักษณะเบ้ ควรประมาณค่า h จากสูตร[17]

$$h_2 = 0.79\,Rn^{-(1/5)}$$

h_2 เป็นความกว้างของช่วง R เป็นค่าพิสัยควอไทล์ (Interquartile Range) หาได้โดยสูตร[18]

---

[16] B.W. Silverman, Density Estimation for Statistics and Data Analysis (London : Chapman and Hall, 1986), p. 45.

[17] Clarke G.M., A Basic Course in Statistics. (London : Edward Arnold), p. 48.

[18] Ibid., p. 48.

$$R = Q_3 - Q_1$$

Q_1 และ Q_3 เป็นค่าควอไทล์ที่ 1 และควอไทล์ที่ 3 ตามลำดับ

จากแนวทางตามที่เสนอมาดังกล่าว และเมื่อพิจารณาประกอบกับประชากรที่มีการแจกแจงปัวส์ซงพารามิเตอร์ λ เท่ากับ 1 2 3 4 5 ผู้วิจัยนำทั้งค่า h_1 และ h_2 มาลองศึกษาดู พร้อมทั้งศึกษาเพิ่มเติมในกรณี h_3 เท่ากับ 0.5 ด้วย การเลือก h_3 เท่ากับ 0.5 ขึ้นมาศึกษาเพราะตัวแปรสุ่มที่มีการแจกแจงปัวส์ซงเป็นตัวแปรสุ่มไม่ต่อเนื่อง ส่วนการประมาณฟังก์ชันการแจกแจงของตัวแปรสุ่มโดยวิธีเคอร์เนล เป็นการประมาณฟังก์ชัน f ซึ่งเป็นฟังก์ชันความหนาแน่นของตัวแปรสุ่มต่อเนื่อง ดังนั้น ผู้วิจัยจึงลองเลือกความกว้างตามแบบที่เคยใช้กันโดยทั่วไป กล่าวคือ

$$P(X = x) = P(x - .5 < X < x + .5)$$

การวิจัยครั้งนี้ผู้วิจัยต้องการศึกษาถึงวิธีการประมาณฟังก์ชันความน่าจะเป็นของประชากร เมื่อประชากรมีการแจกแจงปัวส์ซงพารามิเตอร์ λ เป็น 1 2 3 4 และ 5 ตามลำดับ โดยใช้วิธีของเคอร์เนล และเลือกฟังก์ชันเคอร์เนล K_1 และ K_2 ซึ่งเป็นฟังก์ชันสามเหลี่ยม และฟังก์ชันความหนาแน่นปกติมาตรฐาน ตามลำดับ ส่วนความกว้างของช่วงเลือกศึกษาความกว้าง h_1 h_2 และ h_3 ตามที่กล่าวมาแล้ว เพื่อศึกษาว่าวิธีเคอร์เนลซึ่งเหมาะกับการประมาณฟังก์ชันความหนาแน่นของตัวแปรสุ่มต่อเนื่องจะเหมาะสมหรือไม่ เมื่อนำมาใช้ประมาณฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง