บทที่ 2

ผลงานวิจัยและงานเขียนอื่น ๆ ที่เกี่ยวข้อง

2.1 ผลงานวิจัยที่เกี่ยวข้อง

2.1.1 วิธีการประมาณความหนาแน่น

2.1.1.1 การประมาณฟังก์ชันความหนาแน่นด้วยวิธีเคอร์เนล

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

เมื่อ	$\hat{f}(x)$	คือ	ตัวประมาณฟังก์ชันความหนาแน่นด้วยวิธีเคอร์เนล
	h	คือ	แบนวิดธ์ มีค่าเป็นจำนวนจริงบวก
	n	คือ	ขนาดตัวอย่าง
	$K(u)$	คือ	ฟังก์ชันเคอร์เนล
	x	คือ	ตัวแปรไม่ทราบค่า
	X_i	คือ	ตัวแปรสุ่ม X ตัวที่ i เมื่อ i = 1 , 2 , … , n

คุณสมบัติต่าง ๆ ของการประมาณความหนาแน่นด้วยวิธีเคอร์เนล

1. เมื่อฟังก์ชันเคอร์เนลเป็นฟังก์ชันเคอร์เนลแบบสมมาตร $K(u) = K(-u)$

2. เมื่อฟังก์ชันเคอร์เนลเป็นฟังก์ชันเคอร์เนลแบบสมมาตร จะได้ว่า $\int K(u)du = 1$, $\int uK(u)du = 0$ และ $\int u^2K(u)du > 0$

3. เมื่อฟังก์ชันเคอร์เนลไม่เป็นลบ ณ ที่ใด ๆ และ $\int K(u)du = 1$ แสดงถึงฟังก์ชันเคอร์เนลเป็นฟังก์ชันความหนาแน่นของความน่าจะเป็น ฟังก์ชันความหนาแน่นของความน่าจะเป็นที่ประมาณได้ $(\hat{f}(x))$ จะเป็นฟังก์ชันความหนาแน่นของความน่าจะเป็นด้วย และ $\int \hat{f}(x)dx = 1$

4. ถ้าฟังก์ชันเคอร์เนล $(K(u))$ เป็นฟังก์ชันต่อเนื่องและหาอนุพันธ์ได้

ฟังก์ชันความหนาแน่นของความน่าจะเป็นที่ประมาณได้ $(\hat{f}(x))$ จะเป็นฟังก์ชันต่อเนื่องและหาอนุพันธ์ได้เช่นเดียวกัน

5. ฟังก์ชันความหนาแน่นที่ประมาณได้ $(\hat{f}(x))$ เป็นตัวประมาณที่เอนเอียงของฟังก์ชันความหนาแน่นของความน่าจะเป็น $(f(x))$

6. การกำหนดจุดเริ่มต้น ไม่มีผลต่อการประมาณความหนาแน่นด้วยวิธีเคอร์-เนล

7. สำหรับชุดข้อมูลหนึ่ง ๆ เมื่อกำหนดแบนวิดธ์และกำหนดรูปแบบฟังก์ชัน-เคอร์เนลแล้ว ฟังก์ชันความหนาแน่นแบบเคอร์เนลที่ประมาณขึ้นได้จะมีเพียงหนึ่งฟังก์ชัน

2.1.1.2 **การประมาณฟังก์ชันความหนาแน่นของ** Zhang **และคณะ** (1999)

$$\tilde{f}(x) = \frac{1}{nh}\sum_{i=1}^{n}\left\{K\left(\frac{x-X_i}{h}\right)+K\left(\frac{x+g(X_i)}{h}\right)\right\}$$

เมื่อ $\tilde{f}(x)$ คือ ตัวประมาณฟังก์ชันความหนาแน่นของ Zhang และคณะ

h คือ แบนวิดธ์ มีค่าเป็นจำนวนจริงบวก

n คือ ขนาดตัวอย่าง

$K(u)$ คือ ฟังก์ชันเคอร์เนล

x คือ ตัวแปรไม่ทราบค่า

X_i คือ ตัวแปรสุ่ม X ตัวที่ i เมื่อ i = 1 , 2 , … , n

$g(X_i)$ คือ ฟังก์ชันการแปลง g ของตัวแปรสุ่ม

ฟังก์ชันและตัวแปรต่าง ๆ ที่เกี่ยวข้องในการประมาณฟังก์ชันความหนาแน่นของ Zhang และคณะ (1999) มีรายละเอียดดังนี้

$g(x)$ คือฟังก์ชันการแปลง เป็นฟังก์ชันที่ไม่เป็นลบ มีความต่อเนื่องและเป็นฟังก์ชันเพิ่มเพียงอย่างเดียว (monotonically increasing) จาก $[0,\infty)\rightarrow[0,\infty)$ มีคุณสมบัติคือ

$g^{-1}(0)=0$ นั่นคือ $g(0)=0$

$g'(0)=1$

$g''(0)=[\frac{2f'(0)}{f(0)}]$

ฟังก์ชันการแปลงที่ใช้ในการวิจัยคือ $g(x)=x+dx^2+Ad^2x^3$ เมื่อ A คือค่าคงที่ ซึ่ง $3A>1$ กำหนดค่า $A=0.55$ ดังนั้นจะได้ฟังก์ชันการแปลงคือ $g(x)=x+dx^2+0.55d^2x^3$ (Zhang et al., 1999 , Karunamuni & Alberts, 2005)

d คือ $\frac{f'(0)}{f(0)}$ โดยปกติมักไม่ทราบฟังก์ชันความหนาแน่นของความน่าจะเป็นที่แท้จริง $(f(x))$ จึงไม่สามารถหาค่า d ได้โดยตรง จะต้องประมาณขึ้นโดยใช้ตัวประมาณค่าของ d คือ d_n นั่นคือ

$$d_n=\frac{\log f_n(h)-\log f_n(0)}{h}$$

โดยที่ $f_n^*(h)=\frac{1}{nh}\sum_{i=1}^{n}K(\frac{h-X_i}{h})$

$$f_n(h)=f_n^*(h)+\frac{1}{n^2}$$

$$f_n^*(0)=\frac{1}{nh_0}\sum_{i=1}^{n}K_{(0)}(\frac{-X_i}{h_0})$$

$$f_n(0)=\max(f_n^*(0),\frac{1}{n^2})$$

$$h_0=b(0)h$$

$$b(0)=\left\{\frac{\left[\int_{-\infty}^{\infty}u^2K(u)du\right]^2\int_{-1}^{0}K_0(u)^2du}{\left[\int_{-1}^{0}u^2K_0(u)du\right]^2\int_{-\infty}^{\infty}K(u)^2du}\right\}^{\frac{1}{5}}$$

เมื่อ $K(u)$ คือ ฟังก์ชันเคอร์เนล

$K_{(0)}(u)$ คือ ฟังก์ชันเคอร์เนลของจุดสิ้นสุด (End-point kernel)

จากการกำหนดรูปแบบฟังก์ชันเคอร์เนลและฟังก์ชันเคอร์เนลของจุดสิ้นสุด จะได้ว่า $b(0)=4.427786896$ ดังนั้น $h_0=4.427786896h$ (แสดงวิธีการคำนวณอย่างละเอียดในภาคผนวก ข)

2.1.2 ฟังก์ชันเคอร์เนลที่ใช้ในการวิจัย

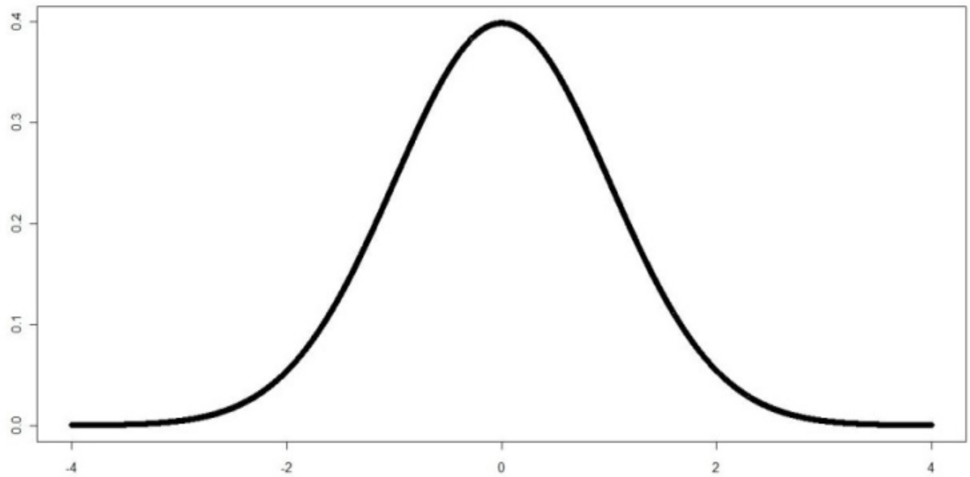
การวิจัยในครั้งนี้มีฟังก์ชันเคอร์เนลที่เกี่ยวข้องทั้งสิ้น 2 รูปแบบ คือ ฟังก์ชันเคอร์เนลแบบเกาส์เซียน (Gaussian kernel) และ ฟังก์ชันเคอร์เนลของจุดสิ้นสุด (End-point kernel)

2.1.2.1 **ฟังก์ชันเคอร์เนลแบบเกาส์เซียน** (Gaussian kernel)

รูปแบบฟังก์ชันเคอร์เนลแบบเกาส์เซียน คือ $K(u) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}u^2)$ รูปร่างฟังก์ชันเคอร์เนลแบบเกาส์เซียน แสดงได้ดังภาพที่ 2.1

ภาพที่ 2.1

รูปร่างฟังก์ชันเคอร์เนลแบบเกาส์เซียน



2.1.2.2 **ฟังก์ชันเคอร์เนลของจุดสิ้นสุด** (End-point kernel)

รูปแบบฟังก์ชันเคอร์เนลของจุดสิ้นสุด คือ $K_{(0)}(u) = 6 + 18u + 12u^2$, $I(-1 \le u \le 0)$

2.1.3 วิธีการเลือกแบนวิดธ์

2.1.3.1 ค่าแบนวิดธ์ที่เหมาะสม (Optimal bandwidth)

ค่าแบนวิดธ์ที่เหมาะสม ถือเป็นแนวคิดพื้นฐานที่สำคัญของการเลือกแบนวิดธ์วิธีต่าง ๆ วิธีการเลือกแบนวิดธ์หลากหลายวิธี ได้อาศัยหลักการของค่าแบนวิดธ์ที่เหมาะสมมาเป็นพื้นฐานของการกำหนดค่าแบนวิดธ์ โดยค่าแบนวิดธ์ที่เหมาะสมคือค่าแบนวิดธ์ที่ทำให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย (Mean Integrated Squared Error : $MISE$) ที่เกิดขึ้นจากการประมาณความหนาแน่นมีค่าน้อยที่สุด ค่าแบนวิดธ์ที่เหมาะสมคือ

$$h_{opt} = \left[\frac{R(K)}{\mu_2(K)^2 R(f'')n} \right]^{\frac{1}{5}}$$

โดยที่ $R(K) = \int K(u)^2 du$

$\mu_2(K) = \int u^2 K(u) du$

$R(f'') = \int f''(x)^2 dx$

เมื่อ h_{opt} คือ ค่าแบนวิดธ์ที่เหมาะสม

n คือ ขนาดตัวอย่าง

$K(u)$ คือ ฟังก์ชันเคอร์เนล

$f(x)$ คือ ฟังก์ชันความหนาแน่นของความน่าจะเป็น

$f''(x)$ คือ อนุพันธ์อันดับที่ 2 ของฟังก์ชันความหนาแน่นของความน่าจะเป็น

จากค่าแบนวิดธ์ที่เหมาะสม พบว่า ค่า $R(K)$ และ $\mu_2(K)$ ขึ้นอยู่กับฟังก์ชันเคอร์เนลซึ่งเป็นส่วนที่ทราบค่า และ $R(f'')$ ขึ้นอยู่กับฟังก์ชันความหนาแน่นของความน่าจะเป็น $(f(x))$ ซึ่งเป็นส่วนที่ไม่ทราบค่า ผลเฉลยของค่าแบนวิดธ์ที่เหมาะสมจะขึ้นอยู่กับฟังก์ชันความหนาแน่นของความน่าจะเป็นที่แท้จริง ดังนั้นวิธีการเลือกแบนวิดธ์หลายวิธีจึงมุ่งเน้นที่จะประมาณค่า $R(f'')$ เพื่อให้ได้ผลลัพธ์ของค่าแบนวิดธ์

2.1.3.2 **วิธี** Direct plug-in (Woodroofe, 1970 , Nadaraya, 1974)

การเลือกแบนวิดธ์ด้วยวิธี Direct plug-in เป็นวิธีที่ทำให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย (Mean Integrated Squared Error : $MISE$) ที่เกิดขึ้นจากการประมาณความหนาแน่นมีค่าน้อยที่สุด แบนวิดธ์ที่ได้จากการเลือกแบนวิดธ์ด้วยวิธี Direct plug-in คือ

$$h_{DPI} = \left[\frac{R(K)}{\mu_2(K)^2 \hat{\psi}_4(g) n} \right]^{\frac{1}{5}}$$

โดยที่ $R(K) = \int K(u)^2 du$

$$\mu_2(K) = \int u^2 K(u) du$$

$$\hat{\psi}_r(g) = n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} L_g^{(r)}(X_i - X_j) \; ; r = 4$$

$$g = \left[\frac{2K^{(4)}(0)}{-\mu_2(K)\psi_6 n} \right]^{\frac{1}{7}}$$

$$\psi_c = \frac{(-1)^{\frac{c}{2}} c!}{(2\hat{\sigma})^{c+1} (\frac{c}{2})! (\pi^{\frac{1}{2}})} \; ; c = 6$$

เมื่อ h_{DPI} คือ แบนวิดธ์ที่ได้จากวิธี Direct plug-in

n คือ ขนาดตัวอย่าง

$K(u)$ คือ ฟังก์ชันเคอร์เนล

g คือ แบนวิดธ์ที่แตกต่างจาก h

$L(u)$ คือ ฟังก์ชันเคอร์เนล อาจเป็นฟังก์ชันเคอร์เนลที่แตกต่างจาก $K(u)$

$\hat{\sigma}$ คือ ค่าประมาณส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม $X_i; i = 1,...,n$

π คือ 3.142857……

$K^{(4)}(u)$ คือ อนุพันธ์อันดับที่ 4 ของฟังก์ชันเคอร์เนล

2.1.3.3 **วิธี** Rules of thumb (Deheuvels, 1977 , Silverman, 1986)

การเลือกแบนวิดธ์ด้วยวิธี Rules of thumb เป็นวิธีการเลือกแบนวิดธ์ที่ง่ายที่สุด ทำได้โดยการแทนค่าส่วนที่ไม่ทราบค่า นั่นคือ $\int f''(x)^2 dx$ ลงในค่าแบนวิดธ์ที่เหมาะสม (Optimal bandwidth)

$$h_{ROT} = 1.06\hat{\sigma}(n^{-\frac{1}{5}})$$

เมื่อ h_{ROT} คือ แบนวิดธ์ที่ได้จากวิธี Rules of thumb

n คือ ขนาดตัวอย่าง

$\hat{\sigma}$ คือ ค่าประมาณส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม $X_i; i = 1,...,n$

2.1.3.4 **วิธี** Least squares cross-validation (Rudemo, 1982 , Bowman, 1984)

การเลือกแบนวิดธ์ด้วยวิธี Least squares cross-validation เป็นวิธีการเลือกแบนวิดธ์ที่มีการศึกษาอย่างกว้างขวางที่สุด (Park & Marron, 1990) เป็นวิธีที่ทำให้ค่าความคลาดเคลื่อนกำลังสองรวม (Integrated Squared Error : ISE) ที่เกิดขึ้นจากการประมาณความหนาแน่นมีค่าน้อยที่สุด ค่าแบนวิดธ์ที่ได้จากวิธี Least squares cross-validation (h_{LSCV}) คือค่าแบนวิดธ์ที่ทำให้ $LSCV(h)$ มีค่าน้อยที่สุด

$$h_{LSCV} = \min_h [LSCV(h)]$$

ค่าความคลาดเคลื่อนกำลังสองรวม คือ

$$ISE(\hat{f}(x)) = \int (\hat{f}(x) - f(x))^2 dx = \int (\hat{f}(x))^2 dx - 2\int \hat{f}(x) f(x) dx + \int f^2(x) dx$$

เมื่อพิจารณา $ISE(\hat{f}(x))$ พบว่า $\int f^2(x)dx$ เป็นส่วนที่มีความเป็นอิสระกับค่าแบนวิดธ์ แต่ $\int (\hat{f}(x))^2 dx$ และ $2\int \hat{f}(x) f(x) dx$ ขึ้นกับค่าแบนวิดธ์ที่เลือกใช้

ฟังก์ชัน Least squares cross-validation กำหนดโดย

$$LSCV(h) = \int (\hat{f}(x))^2 dx - 2\int \hat{f}(x) f(x) dx = \frac{1}{n}\sum_{i=1}^{n}\int (\hat{f}_{-i}(x))^2 dx - \frac{2}{n}\sum_{i=1}^{n} \hat{f}_{-i}(X_i)$$

เมื่อ h_{LSCV} คือ แบนวิดธ์ที่ได้จากวิธี Least squares cross-validation

n คือ ขนาดตัวอย่าง

$\hat{f}(x)$ คือ ตัวประมาณของฟังก์ชันความหนาแน่นของความน่าจะเป็น

$f(x)$ คือ ฟังก์ชันความหนาแน่นของความน่าจะเป็น

$\hat{f}_{-i}(x)$ คือ ตัวประมาณของฟังก์ชันความหนาแน่นของความน่าจะเป็น โดยปราศจากตัวแปรสุ่ม $X_i; i = 1,...,n$ (leave one out)

2.1.3.5 **วิธี** Silverman’s rules of thumb (Silverman, 1986)

การเลือกแบนวิดธ์ด้วยวิธี Silverman’s rules of thumb เป็นวิธีการเลือกแบนวิดธ์ที่ถูกปรับปรุงมาจากการเลือกแบนวิดธ์ด้วยวิธี Rules of thumb

$$h_{SROT} = 0.9A(n^{-\frac{1}{5}})$$

เมื่อ h_{SROT} คือ แบนวิดธ์ที่ได้จากวิธี Silverman’s Rules of thumb

n คือ ขนาดตัวอย่าง

A คือ ค่าที่น้อยที่สุดระหว่างค่าประมาณส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม $X_i; i = 1,...,n$ กับ $\frac{IQR}{1.34} = \min[\hat{\sigma}, \frac{IQR}{1.34}]$

IQR คือ พิสัยระหว่างควอร์ไทล์ (Interquartile range)

$\hat{\sigma}$ คือ ค่าประมาณส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม $X_i; i = 1,...,n$

2.1.3.6 **วิธี** Biased cross-validation (Scott & Terrell, 1987)

การเลือกแบนวิดธ์ด้วยวิธี Biased cross-validation เป็นวิธีที่ทำให้ค่า Asymptotic Mean Integrated Squared Error $(AMISE)$ ที่เกิดขึ้นจากการประมาณความหนาแน่นมีค่าน้อยที่สุด ค่าแบนวิดธ์ที่ได้จากวิธี Biased cross-validation (h_{BCV}) คือค่าแบนวิดธ์ที่ทำให้ $BCV(h)$ มีค่าน้อยที่สุด

$$h_{BCV} = \min_{h}[BCV(h)]$$

ฟังก์ชัน Biased cross-validation กำหนดโดย

$$BCV(h) = \frac{1}{nh}\int K(u)^2 du + \frac{h^4}{4}[\int u^2 K(u)du]^2[\int \hat{f}''(x)^2 dx - \frac{1}{nh^5}\int K''(u)^2 du]$$

เมื่อ h_{BCV} คือ แบนวิดธ์ที่ได้จากวิธี Biased cross-validation

n คือ ขนาดตัวอย่าง

$K(u)$ คือ ฟังก์ชันเคอร์เนล

$\hat{f}(x)$ คือ ตัวประมาณของฟังก์ชันความหนาแน่นของความน่าจะเป็น

$\hat{f}''(x)$ คือ อนุพันธ์อันดับที่ 2 ของตัวประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็น

2.1.3.7 **วิธี** Solve the equation plug-in (Sheather & Jones, 1991)

การเลือกแบนวิดธ์ด้วยวิธี Solve the equation plug-in เป็นวิธีที่ทำให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย (Mean Integrated Squared Error : $MISE$) ที่เกิดขึ้นจากการประมาณความหนาแน่นมีค่าน้อยที่สุด แบนวิดธ์ที่ได้จากการเลือกแบนวิดธ์ด้วยวิธี Solve the equation plug-in คือ ผลเฉลยของสมการ นั่นคือ

$$h_{STE} = \left[\frac{R(K)}{\mu_2(K)^2 \hat{S}_D(\hat{\alpha}_2(h_{STE}))n} \right]^{\frac{1}{5}}$$

โดยที่ $R(K) = \int K(u)^2 du$

$\mu_2(K) = \int u^2 K(u) du$

$\hat{S}_D(\alpha) = \{n(n-1)\}^{-1} \alpha^{-5} \sum_{i=1}^{n} \sum_{j=1}^{n} \phi^{(4)} \{\frac{(X_i - X_j)}{\alpha}\}$

$\hat{\alpha}_2(h) = 1.357 \left\{ \frac{\hat{S}_D(a)}{\hat{T}_D(b)} \right\}^{\frac{1}{7}} h^{\frac{5}{7}}$

$\hat{T}_D(b) = -\{n(n-1)\}^{-1} b^{-7} \sum_{i=1}^{n} \sum_{j=1}^{n} \phi^{(6)} \{\frac{(X_i - X_j)}{b}\}$

$a = 0.920(IQR)(n)^{-\frac{1}{7}}$

$b = 0.912(IQR)(n)^{-\frac{1}{9}}$

เมื่อ h_{STE} คือ แบนวิดธ์ที่ได้จากวิธี Solve the equation plug-in

n คือ ขนาดตัวอย่าง

$K(u)$ คือ ฟังก์ชันเคอร์เนล

$\phi(u)$ คือ ฟังก์ชันเคอร์เนล อาจเป็นฟังก์ชันเคอร์เนลที่แตกต่างจาก $K(u)$

IQR คือ พิสัยระหว่างควอร์ไทล์ (Interquartile range)

2.1.4 รูปแบบฟังก์ชันความหนาแน่นของความน่าจะเป็นที่ใช้ในการวิจัย

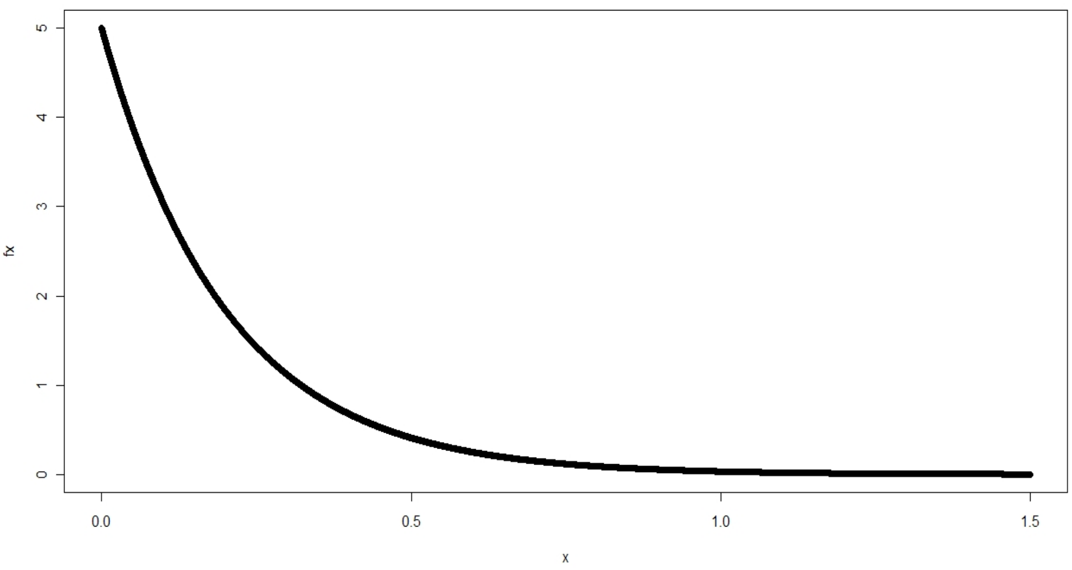
การวิจัยในครั้งนี้ทำการศึกษาฟังก์ชันความหนาแน่นของความน่าจะเป็นจำนวน 2 รูปแบบ ทั้ง 2 รูปแบบล้วนเป็นฟังก์ชันยกกำลัง (Exponential function) ได้แก่ $f(x)=5e^{-5x}$ เมื่อ $x>0$ และ $f(x)=\frac{5}{4}(1+15x)e^{-5x}$ เมื่อ $x>0$

2.1.4.1 $f(x)=5e^{-5x}$ **เมื่อ** $x>0$

ลักษณะกราฟฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1 แสดงได้-ดังภาพที่ 2.2

ภาพที่ 2.2

ฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 1

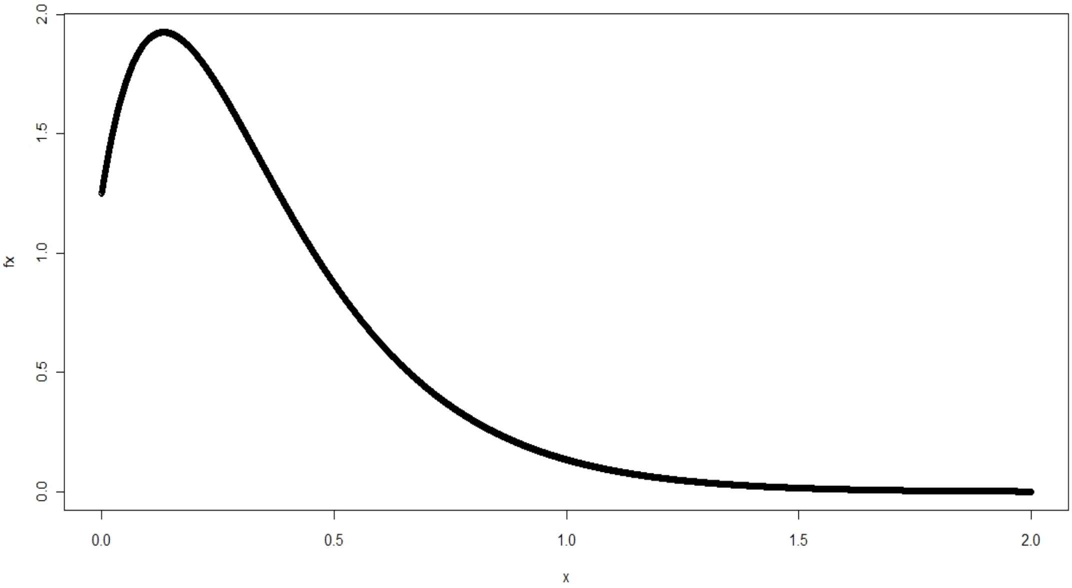


2.1.4.2 $f(x) = \frac{5}{4}(1+15x)e^{-5x}$ **เมื่อ** $x > 0$

ลักษณะกราฟฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2 แสดงได้-ดังภาพที่ 2.3

ภาพที่ 2.3

ฟังก์ชันความหนาแน่นของความน่าจะเป็นรูปแบบที่ 2



2.1.5 เกณฑ์ในการพิจารณา

ในการประมาณความหนาแน่นด้วยวิธีเคอร์เนล จำเป็นที่จะต้องมีเครื่องมือที่จะใช้ในการชี้วัดความแตกต่างระหว่างฟังก์ชันความหนาแน่นของความน่าจะเป็นที่ประมาณได้ กับฟังก์ชันความหนาแน่นของความน่าจะเป็นที่แท้จริง เกณฑ์ในการชี้วัดที่นิยมใช้อย่างกว้างขวางที่สุดและสามารถชี้วัดความแม่นยำโดยรวม (global accuracy) ได้เป็นอย่างดีคือค่าความคลาดเคลื่อน-กำลังสองรวมเฉลี่ย (Mean Integrated Squared Error : $MISE$) (Silverman, 1986, pp. 35-36)

$$MISE(\hat{f}(x)) = E\int \left[\hat{f}(x) - f(x) \right]^2 dx$$

เมื่อ $MISE$ คือ ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย

$\hat{f}(x)$ คือ ตัวประมาณของฟังก์ชันความหนาแน่นของความน่าจะเป็น

$f(x)$ คือ ฟังก์ชันความหนาแน่นของความน่าจะเป็น

2.2 งานเขียนอื่น ๆ ที่เกี่ยวข้อง

ในปีค.ศ. 1956 Rosenblatt นำเสนอแนวทางการประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบไม่อิงพารามิเตอร์สมัยใหม่ ได้เสนอวิธีการประมาณแบบเคอร์เนลขึ้นเป็นครั้งแรก รวมทั้งพิสูจน์ให้เห็นถึงความเอนเอียงของตัวประมาณฟังก์ชันความหนาแน่นของความน่าจะเป็น และกล่าวว่าสำหรับการประมาณความหนาแน่นจะมีความคลาดเคลื่อนเกิดขึ้นเสมอ จึงจำเป็นที่จะต้องมีเครื่องมือที่จะใช้ในการชี้วัดความแตกต่างระหว่างฟังก์ชันความหนาแน่นของความน่าจะเป็นที่ประมาณขึ้น กับฟังก์ชันความหนาแน่นของความน่าจะเป็นที่แท้จริง ได้เสนอให้ใช้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย (Mean Integrated Squared Error : $MISE$) เพื่อชี้วัดความคลาดเคลื่อนที่เกิดขึ้นจากการประมาณความหนาแน่น ต่อมา Hodges และ Lehmann (1956) เสนอให้ใช้ฟังก์ชันเคอร์เนลแบบ Epanechnikov เพื่อทำให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยในการประมาณความหนาแน่นมีค่าน้อยที่สุด และในปีค.ศ. 1986 Silverman เปรียบเทียบประสิทธิภาพของฟังก์ชันเคอร์เนลแบบ Epanechnikov กับ ฟังก์ชันเคอร์เนลรูปแบบต่าง ๆ พบว่าฟังก์ชันเคอร์เนลแบบ Epanechnikov มีประสิทธิภาพสูงสุด ในปีค.ศ. 1996 Simonoff กล่าวถึงประสิทธิภาพของฟังก์ชันเคอร์เนลรูปแบบต่าง ๆ ว่า ไม่ว่าจะเลือกใช้ฟังก์ชันเคอร์เนลรูปแบบใดก็ตาม จะไม่ส่งผลต่อค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยมากนัก เพราะฟังก์ชันเคอร์เนลแต่ละรูปแบบนั้นมีประสิทธิภาพไม่แตกต่างกันอย่างชัดเจน (Simonoff, 1996, pp. 44) ในปีค.ศ. 2004 Sheather กล่าวว่า ฟังก์ชันเคอร์เนลแบบเกาส์เซียนถือเป็นรูปแบบฟังก์ชันเคอร์เนลที่ได้รับความนิยมอย่างแพร่หลาย

ในปีค.ศ. 1990 Hardle พิจารณาลักษณะของการประมาณความหนาแน่นในสองสถานการณ์ คือ เมื่อกำหนดค่าแบนวิดธ์เท่ากันและใช้ฟังก์ชันเคอร์เนลรูปแบบต่างกัน พบว่าลักษณะกราฟฟังก์ชันความหนาแน่นของความน่าจะเป็นมีความแตกต่างกันไม่มาก และเมื่อกำหนดค่าแบนวิดธ์แตกต่างกันและใช้ฟังก์ชันเคอร์เนลรูปแบบเดียวกัน พบว่าลักษณะกราฟฟังก์ชันความหนาแน่นของความน่าจะเป็นมีความแตกต่างกันมาก จากทั้ง 2 สถานการณ์จึงสรุปได้ว่า การกำหนดฟังก์ชันเคอร์เนลมีอิทธิพลต่อการประมาณความหนาแน่นไม่มากนัก แต่การกำหนดค่าแบนวิดธ์มีอิทธิพลต่อการประมาณความหนาแน่นมากกว่า ข้อสรุปดังกล่าวได้รับการสนับสนุนโดย Mugdadi และ Ahmad ที่กล่าวไว้ในปีค.ศ. 2004 ว่า ทั้งการวิเคราะห์ทางทฤษฎีและการจำลองข้อมูล พบว่า ในกรณีที่ตัวแปรสุ่มเป็นอิสระต่อกันและมีการแจกแจงเดียวกัน การกำหนดรูปแบบฟังก์ชันเคอร์เนลไม่ใช่สิ่งที่มีผลต่อการประมาณความหนาแน่นมากนัก สิ่งที่สำคัญที่สุดในการประมาณความหนาแน่น คือการเลือกแบนวิดธ์ให้มีความเหมาะสม

เนื่องจากการเลือกแบนวิดธ์ถือเป็นสิ่งสำคัญในการประมาณฟังก์ชันความหนาแน่น-ของความน่าจะเป็นด้วยวิธีเคอร์เนล ทำให้มีนักวิจัยจำนวนมากนำเสนอวิธีการเลือกแบนวิดธ์ โดยในปีค.ศ. 1970 การเลือกแบนวิดธ์ด้วยวิธี Direct plug-in ถูกเสนอขึ้นโดย Woodroofe เป็นการเลือกแบนวิดธ์บนพื้นฐานของการมีค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย (Mean Integrated Squared Error : $MISE$) น้อยที่สุด ต่อมาในปีค.ศ. 1977 การเลือกแบนวิดธ์ด้วยวิธี Rules of thumb ถูกเสนอขึ้นโดย Deheuvels ถือเป็นวิธีที่ง่ายที่สุดในการเลือกค่าแบนวิดธ์ ทำได้โดยการแทนค่า $\int f''(x)^2 dx$ ซึ่งเป็นส่วนที่ไม่ทราบค่าลงในค่าแบนวิดธ์ที่เหมาะสม (Optimal bandwidth) และในปีค.ศ. 1986 Silverman พิจารณาการเลือกแบนวิดธ์ด้วยวิธี Rules of thumb กำหนดให้ฟังก์ชันความหนาแน่นของความน่าจะเป็นที่แท้จริง $(f(x))$ เป็นการแจกแจงแบบปกติ-มาตรฐาน (Standard normal distribution) กำหนดรูปแบบฟังก์ชันเคอร์เนลแบบเกาส์เซียน จะได้ค่าแบนวิดธ์คือ $h = 1.06\hat{\sigma}(n^{-\frac{1}{5}})$ ซึ่งใช้ได้ดีเมื่อการแจกแจงที่แท้จริงเป็นการแจกแจงแบบปกติ แต่เมื่อการแจกแจงที่แท้จริงมีฐานนิยมหลายค่า การใช้แบนวิดธ์ค่านี้ในการประมาณความหนาแน่นอาจทำได้ไม่ดีนัก จึงเสนอให้ปรับลดค่าคงที่จาก 1.06 เหลือ 0.9 คูณกับค่าที่น้อยที่สุดระหว่างส่วนเบี่ยงเบนมาตรฐานตัวอย่าง $(\hat{\sigma})$ กับพิสัยระหว่างควอร์ไทล์หารด้วย 1.34 $\left(\frac{IQR}{1.34}\right)$ จากนั้นนำไปคูณกับ $n^{-\frac{1}{5}}$ ส่งผลให้เกิดอีกทางเลือกหนึ่งของค่าแบนวิดธ์ เรียกว่าวิธี Silverman's rules of thumb ซึ่งสามารถใช้ได้ทั้งในกรณีที่ตัวอย่างสุ่มมีการแจกแจงที่แท้จริงแบบฐานนิยมค่าเดียวและแบบฐานนิยมสองค่า จะทำให้ความเรียบของกราฟความหนาแน่นมีความสมดุลยิ่งขึ้น (Silverman, 1986, pp. 46-48) ในปีค.ศ. 1984 Bowman ได้เสนอการเลือกแบนวิดธ์ด้วยวิธี Least squares cross-validation เป็นการเลือกแบนวิดธ์บนพื้นฐานของการมีค่าความ-คลาดเคลื่อนกำลังสองรวม (Integrated Squared Error : ISE) น้อยที่สุด ต่อมาในปีค.ศ. 1987 Scott และ Terrell ได้เสนอการเลือกแบนวิดธ์ด้วยวิธี Biased cross-validation เป็นการเลือกแบนวิดธ์บนพื้นฐานของการมีค่า Asymptotic Mean Integrated Squared Error $(AMISE)$ น้อยที่สุด ในปีค.ศ. 1991 Sheather และ Jones ได้เสนอการเลือกแบนวิดธ์ด้วยวิธี Solve the equation plug-in ซึ่งถูกพัฒนามาจากวิธีของ Park และ Marron (1990) เป็นการเลือกแบนวิดธ์บนพื้นฐานของการมีค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ย (Mean Integrated Squared Error : $MISE$) น้อยที่สุด

ในปีค.ศ. 1990 Park และ Marron นำเสนอวิธีการเลือกแบนวิดธ์โดยพัฒนามาจากการเลือกแบนวิดธ์ด้วยวิธี Direct Plug-in เรียกว่า การเลือกแบนวิดธ์ด้วยวิธีของ Park และ

Marron และทำการศึกษาเปรียบเทียบวิธีการเลือกแบนวิดธ์ 4 วิธี คือวิธี Oversmooth bandwidth วิธีของ Park และ Marron วิธี Least squares cross-validation และ วิธี Biased cross-validation ใช้รูปแบบการแจกแจง 4 รูปแบบ คือ การแจกแจงแบบปกติมาตรฐาน การแจกแจงแบบปกติผสมที่มีค่าเฉลี่ยแตกต่างกัน การแจกแจงแบบปกติผสมที่มีค่าผิดปกติ การแจกแจงแบบปกติผสมที่มีความแปรปรวนแตกต่างกัน พบว่าวิธีของ Park และ Marron ให้ผลลัพธ์ที่ดีในเกือบทุกสถานการณ์ และให้ผลลัพธ์ดีที่สุดเมื่อตัวแปรสุ่มเกิดจากการแจกแจงแบบปกติผสมที่มีค่าเฉลี่ยแตกต่างกันและการแจกแจงแบบปกติผสมที่มีความแปรปรวนแตกต่างกัน ต่อมาในปีค.ศ. 1996 Jones , Marron และ Sheather ศึกษาเปรียบเทียบวิธีการเลือกแบนวิดธ์ โดยแบ่งวิธีการเลือกแบนวิดธ์ออกเป็น 2 ยุค วิธีการเลือกแบนวิดธ์ในยุคแรกถูกนำเสนอขึ้นก่อนปี ค.ศ. 1990 ได้แก่ วิธี Rules of thumb วิธี Least squares cross-validation และ วิธี Biased cross-validation วิธีการเลือกแบนวิดธ์ในยุคที่สองถูกนำเสนอขึ้นหลังปีค.ศ. 1990 ได้แก่ วิธี Solve the equation plug-in และ วิธี Smoothed Bootstrap ทำการเปรียบเทียบวิธีการเลือกแบนวิดธ์จำนวน 4 วิธี คือ วิธี Rules of thumb วิธี Least squares cross-validation วิธี Biased cross-validation และ วิธี Solve the equation plug-in โดยใช้ข้อมูลจริงคือข้อมูลมวลกล้ามเนื้อมนุษย์ ซึ่งมีการแจกแจงที่แท้จริงแบบฐานนิยมสองค่า (bimodal) จากการวิจัยพบว่า การใช้แบนวิดธ์จากวิธี Rules of thumb มาประมาณความหนาแน่น จะได้ลักษณะกราฟความหนาแน่นที่ค่อนข้างเรียบมากกว่าปกติ ส่งผลให้ไม่สามารถเข้าถึงการมีฐานนิยมสองค่าหรือพลาดลักษณะที่สำคัญของข้อมูล เนื่องจากแบนวิดธ์มีค่ามากเกินไป แบนวิดธ์ที่ได้จากวิธี Least squares cross-validation ทำให้กราฟความหนาแน่นมีความเรียบน้อยกว่าปกติอย่างชัดเจน ส่งผลให้ผู้พบเห็นไม่สามารถที่จะเข้าใจลักษณะของข้อมูล เนื่องจากลักษณะกราฟความหนาแน่นไม่สามารถสะท้อนลักษณะที่แท้จริงของข้อมูลออกมาได้ แบนวิดธ์ที่ได้จากวิธี Biased cross-validation ทำให้กราฟความหนาแน่นที่ประมาณได้ มีความเรียบมากกว่าปกติและมีความเรียบมากกว่าการใช้แบนวิดธ์ที่ได้จากวิธี Rules of thumb ส่งผลให้พลาดการเข้าถึงการมีฐานนิยมสองค่าอย่างชัดเจน แบนวิดธ์ที่ได้จากวิธี Solve the equation plug-in สามารถประมาณความหนาแน่นได้ดีที่สุดเมื่อเทียบกับ 3 วิธีข้างต้น และแสดงให้เห็นลักษณะกราฟความหนาแน่นที่แสดงถึงการมีฐานนิยมสองค่าได้ดี จากผลการวิจัยสรุปได้ว่า สำหรับข้อมูลจริงชุดดังกล่าว วิธีการเลือกแบนวิดธ์ในยุคที่สองเหมาะสมที่จะใช้ในการเลือกแบนวิดธ์และเหมาะสำหรับวิเคราะห์ด้วยคอมพิวเตอร์

ในปีค.ศ. 1998 Zhang และ Karunamuni เสนอรูปแบบฟังก์ชันเคอร์เนลของจุดสิ้นสุด (End-point kernel) และศึกษาถึงปัญหาการประมาณความหนาแน่น ณ จุดสิ้นสุดของ

ฟังก์ชันความหนาแน่นของความน่าจะเป็นที่สิ้นสุดที่จุดศูนย์ โดยประมาณความหนาแน่นด้วยวิธี The boundary kernel method เพื่อแก้ไขปัญหาการประมาณความหนาแน่นบริเวณขอบเขต และแสดงให้เห็นว่า การใช้ $K_{(0)}(u)$ คือ ฟังก์ชันเคอร์เนลของจุดสิ้นสุด (End-point kernel) เป็นฟังก์ชันเคอร์เนลที่เหมาะสมสำหรับจุดสิ้นสุด เนื่องจากในการประมาณความหนาแน่น ณ จุดสิ้นสุด จะทำให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) มีค่าน้อยที่สุด ต่อมาในปีค.ศ. 1999 Zhang , Karunamuni และ Jones เสนอวิธีการประมาณความหนาแน่นแบบใหม่เพื่อใช้ในการแก้ปัญหาการประมาณความหนาแน่นบริเวณขอบเขต เรียกว่า วิธีการประมาณความหนาแน่นของ Zhang และคณะ (Zhang, Karunamuni, & Jones, 1999) คุณสมบัติที่ดีของวิธีดังกล่าว คือ ตัวประมาณความหนาแน่นจะไม่เป็นลบอย่างแน่นอน Zhang และคณะ ทำการศึกษาเปรียบเทียบวิธีการประมาณความหนาแน่นที่เสนอขึ้น กับวิธีอื่น ๆ อีก 5 วิธี ได้แก่ The boundary kernel method (Zhang & Karunamuni, 1998) , The local linear method , The local linear method with bandwidth variation , The pseudodata method (Cowling & Hall, 1996) และ The nonnegative adaptation estimator (Jones & Foster, 1996) โดยใช้ฟังก์ชันความ-หนาแน่นของความน่าจะเป็น 3 รูปแบบ ในการวิจัยใช้ค่าแบนวิดธ์ที่เหมาะสม ใช้ฟังก์ชันเคอร์-เนลแบบ Epanechnikov และขนาดตัวอย่างเท่ากับ 200 จากผลการศึกษาเปรียบเทียบ พบว่า วิธีการประมาณความหนาแน่นของ Zhang และคณะ ให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยน้อยที่สุดสำหรับฟังก์ชันความหนาแน่นของความน่าจะเป็น 2 รูปแบบ โดยทั้งสองรูปแบบล้วนเป็นฟังก์ชันยกกำลัง (Exponential function) และในปีค.ศ. 2005 Karunamuni และ Alberts เสนอวิธีการประมาณความหนาแน่นแบบใหม่ที่พัฒนาจากวิธีการประมาณความหนาแน่นของ Zhang และคณะ (1999) เพื่อใช้ในการแก้ปัญหาการประมาณความหนาแน่นบริเวณขอบเขต คือวิธีของ Karunamuni และ Alberts ทำการศึกษาเปรียบเทียบวิธีการประมาณความหนาแน่นที่เสนอขึ้น กับวิธีอื่น ๆ อีก 5 วิธี คือ The boundary kernel with bandwidth variation (Zhang & Karunamuni, 1998) , The local linear method , วิธีการประมาณความหนาแน่นของ Zhang และคณะ (Zhang, Karunamuni, & Jones, 1999) , The nonnegative adaptation estimator (Jones & Foster, 1996) และ The method of Hall and Park (Hall & Park, 2002) ใช้ฟังก์ชันความ-หนาแน่นของความน่าจะเป็น 4 รูปแบบ ในการวิจัยใช้ค่าแบนวิดธ์ที่เหมาะสม ใช้ฟังก์ชันเคอร์-เนลแบบ Epanechnikov และขนาดตัวอย่างเท่ากับ 200 จากผลการศึกษาเปรียบเทียบ พบว่า วิธีการประมาณความหนาแน่นของ Karunamuni และ Alberts ให้ค่าความคลาดเคลื่อนกำลัง-สองรวมเฉลี่ยน้อยที่สุดสำหรับฟังก์ชันความหนาแน่นของความน่าจะเป็น 1 รูปแบบ ส่วนวิธีการ

ประมาณความหนาแน่นของ Zhang และคณะ ให้ค่าความคลาดเคลื่อนกำลังสองรวมเฉลี่ยน้อยที่สุดสำหรับฟังก์ชันความหนาแน่นของความน่าจะเป็น 2 รูปแบบ โดยทั้งสองรูปแบบล้วนเป็นฟังก์ชันยกกำลัง (Exponential function)