Conditional Distributions and Function of Random Variables

รศ.ดร. วีระชาติ กิเลนทอง
มหาวิทยาลัยหอการค้าไทย

การแจกแจงแบบมีเงื่อนไข (Conditional Distributions)

- การวิเคราะห์ทางการเงินและเศรษฐศาสตร์มักให้ความสนใจในการพยากรณ์โดยมี “**เงื่อนไขมาจากสิ่งที่เราทราบ**” แล้ว
  - ต้องการทราบว่า ผลตอบแทนของหลักทรัพย์มีค่าเท่าไหร่ หากเราทราบว่า “<u>ธนาคารแห่งประเทศไทยประกาศขึ้นอัตราดอกเบี้ยนโยบาย</u>”
  - ต้องการทราบว่า อัตราการออมของครัวเรือนในชนบท “<u>ที่ระดับรายได้ต่างๆ</u>” มีค่าเป็นเท่าใด
- รูปแบบเหล่านี้คือ “**รูปแบบของการแจกแจงแบบมีเงื่อนไข (conditional distribution)**”

การแจกแจงแบบมีเงื่อนไขของตัวแปรสุ่มไม่ต่อเนื่อง

- การแจกแจงแบบมีเงื่อนไขของตัวแปรสุ่มไม่ต่อเนื่อง (discrete) “**เป็นผลโดยตรง**” จากหลักการความน่าจะเป็นแบบมีเงื่อนไข (conditional probability)
- ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ภายใต้เงื่อนไขของตัวแปรสุ่ม $Y = y$ นิยามได้เป็น

$$g_1(x|y) \equiv Pr(X = x|Y = y) = \frac{Pr(X = x \text{ และ } Y = y)}{Pr(Y = y)} = \frac{f(x, y)}{f_2(y)} \qquad (1)$$

นิยาม: การแจกแจงแบบมีเงื่อนไขของตัวแปรสุ่มไม่ต่อเนื่อง

Definition

กำหนดให้ X และ Y คือตัวแปรสุ่มไม่ต่อเนื่อง (discrete) ที่มี $f(x,y)$ แทนฟังก์ชันความน่าจะเป็นร่วม (joint p.f.) และ $f_2(y)$ แทนฟังก์ชันความน่าจะเป็นตามขอบ (marginal p.f.) ของตัวแปรสุ่ม Y ดังนั้น สำหรับค่าจำนวนจริง y ที่ $f_2(y) > 0$

$$g_1(x|y) \equiv \frac{f(x,y)}{f_2(y)} \tag{2}$$

เรามักเรียก g_1 ว่าฟังก์ชันความน่าจะเป็นแบบมีเงื่อนไข (conditional p.f.) ของ X ภายใต้เงื่อนไขของ Y ส่วนฟังก์ชัน $g_1(\cdot|y)$ คือการแจกแจงแบบมีเงื่อนไข (conditional distribution) ของ X เมื่อค่าของ $Y = y$ ในทำนองเดียวกัน การแจกแจงแบบมีเงื่อนไข (conditional distribution) ของ Y เมื่อค่าของ $X = x$ คือ

$$g_2(y|x) \equiv \frac{f(x,y)}{f_1(x)} \tag{3}$$

สำหรับค่าจำนวนจริง x ที่ $f_1(x) > 0$

นิยาม: การแจกแจงแบบมีเงื่อนไขของตัวแปรสุ่มต่อเนื่อง

Definition

กำหนดให้ X และ Y คือตัวแปรสุ่มต่อเนื่อง (continuous) ที่มี $f(x,y)$ แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) และ $f_1(x)$ และ $f_2(y)$ แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของตัวแปรสุ่ม X และ Y ตามลำดับ ดังนั้น สำหรับค่าจำนวนจริง y ที่ $f_2(y) > 0$ การแจกแจงแบบมีเงื่อนไข (conditional distribution) ของ X เมื่อค่าของ $Y = y$

$$g_1(x|y) \equiv \frac{f(x,y)}{f_2(y)}, \text{ สำหรับ } -\infty < x < \infty \tag{4}$$

ในทำนองเดียวกัน สำหรับค่าจำนวนจริง x ที่ $f_1(x) > 0$ การแจกแจงแบบมีเงื่อนไข (conditional distribution) ของ Y เมื่อค่าของ $X = x$ คือ

$$g_2(y|x) \equiv \frac{f(x,y)}{f_1(x)}, \text{ สำหรับ } -\infty < x < \infty \tag{5}$$

ส่วนกรณีที่ $f_2(y) = 0$ หรือ $f_1(x) = 0$ เราสามารถกำหนดค่าของการแจกแจงแบบมีเงื่อนไข $g_1(x|y)$ หรือ $g_2(y|x)$ ตามลำดับ ได้โดยอิสระ ตราบเท่าที่การแจกแจงแบบมีเงื่อนไขนั้นเป็นฟังก์ชันความหนาแน่นของความน่าจะเป็น

ทฤษฎี: การแจกแจงแบบมีเงื่อนไขเป็นการแจกแจงแบบหนึ่ง

Theorem

กำหนดให้ X และ Y คือตัวแปรสุ่มต่อเนื่อง (continuous) สำหรับค่า y ใดๆ การแจกแจงแบบมีเงื่อนไข (conditional distribution) $g_1(x|y)$ มีคุณสมบัติเป็นฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ที่เป็นฟังก์ชันของตัวแปร x ในทำนองเดียวกัน สำหรับค่า x ใดๆ การแจกแจงแบบมีเงื่อนไข (conditional distribution) $g_2(y|x)$ มีคุณสมบัติเป็นฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ที่เป็นฟังก์ชันของตัวแปร y

การพิสูจน์: การแจกแจงแบบมีเงื่อนไขเป็นการแจกแจงแบบหนึ่ง

Proof.

- เราจะแสดงการพิสูจน์สำหรับ $g_2\left(y|x\right)$ เท่านั้น โดยเริ่มจากกรณีที่ $f_1\left(x\right) > 0$
  - การแจกแจงแบบมีเงื่อนไข (conditional distribution) $g_2\left(y|x\right)$ จะต้องสอดคล้องตามสมการที่ 5 ซึ่งมีผลทำให้ $g_2\left(y|x\right) \geq 0$
- เนื่องจากทั้งตัวตั้งมีค่าเป็นบวกหรือศูนย์ส่วนตัวหารมีค่ามากกว่าศูนย์ ส่วนที่เหลือคือการพิสูจน์ว่า $g_2\left(y|x\right)$ สอดคล้องกับ "**คุณสมบัติผลรวมเท่ากับหนึ่ง**"

$$\int_{-\infty}^{\infty} g_2\left(y|x\right) dy = \int_{-\infty}^{\infty} \frac{f\left(x,y\right)}{f_1\left(x\right)} dy = \frac{1}{f_1\left(x\right)} \int_{-\infty}^{\infty} f\left(x,y\right) dy = \frac{1}{f_1\left(x\right)} f_1\left(x\right) = 1$$

- ความน่าจะเป็นตามขอบ (marginal p.d.f.) ของ ตัวแปรสุ่ม X คือ $f_1\left(x\right) = \int_{-\infty}^{\infty} f\left(x,y\right) dy$
- กรณีที่ $f_1\left(x\right) = 0$ นิยามข้างต้น ได้กำหนดให้ $g_2\left(y|x\right)$ มีคุณสมบัติเป็นฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) อยู่แล้ว
- กรณีของตัวแปรสุ่มไม่ต่อเนื่อง (discrete) การแจกแจงแบบมีเงื่อนไขมีคุณสมบัติเป็นฟังก์ชันความน่าจะเป็น (p.f.) เช่นเดียวกัน

ความสัมพันธ์ระหว่างการแจกแจงแบบมีเงื่อนไขกับการแจกแจงร่วม

Theorem

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่มีการแจกแจง $f_1(x)$ *และ* $f_2(y)$ *ตามลำดับ ในทำนองเดียวกัน กำหนดให้* $g_1(x|y)$ *และ* $g_2(y|x)$ *แทนการแจกแจงแบบมีเงื่อนไขของ X เมื่อค่าของ* $Y = y$ *และ การแจกแจงแบบมีเงื่อนไขของ Y เมื่อค่าของ* $X = x$ *ตามลำดับ ดังนั้น การแจกแจงร่วมของ X และ Y สามารถหาค่าได้จาก*

$$f(x,y) = g_1(x|y) f_2(y), \tag{6}$$

$$f(x,y) = g_2(y|x) f_1(x) \tag{7}$$

โดยที่สมการแรกเป็นจริงสำหรับทุกๆ ค่า x และค่า y ที่ $f_2(y) > 0$ *และในทางกลับกันสมการที่สองเป็นจริงสำหรับทุกๆ ค่า y และค่า x ที่* $f_1(x) > 0$

กฎของความน่าจะเป็นรวม (Law of total probability)

Theorem

ฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) และฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบมีเงื่อนไข (conditional p.d.f.) สำหรับตัวแปรสุ่ม X และ Y สอดคล้องกับกฎของความน่าจะเป็นรวม (Law of total probability) ดังต่อไปนี้

$$f_1(x) = \int_{-\infty}^{\infty} g_1(x|y) f_2(y)\, dy, \tag{8}$$

$$f_2(y) = \int_{-\infty}^{\infty} g_2(y|x) f_1(x)\, dx \tag{9}$$

โดยที่ $f_1(x)$ และ $f_2(y)$ แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของ X และ Y ตามลำดับ และ $g_1(x|y)$ แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบมีเงื่อนไขของ X เมื่อค่าของ $Y = y$ และ $g_2(y|x)$ แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบมีเงื่อนไขของ Y เมื่อค่าของ $X = x$

ทฤษฎีบทของเบส์ (Bayes' Theorem)

Theorem (ทฤษฎีบทของเบส์ (Bayes' Theorem))

กำหนดให้ X และ Y เป็นตัวแปรสุ่ม โดยที่ $f_1(x)$ และ $f_2(y)$ แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของ X และ Y ตามลำดับ ความสัมพันธ์ระหว่างฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบมีเงื่อนไข (conditional p.d.f.) สามารถเขียนได้ในรูปสมการดังต่อไปนี้

$$g_1(x|y) = \frac{g_2(y|x) f_1(x)}{f_2(y)} \tag{10}$$

$$g_2(y|x) = \frac{g_1(x|y) f_2(y)}{f_1(x)} \tag{11}$$

- ในทางปฏิบัติทฤษฎีของเบส์ (Bayes' Theorem) ช่วยให้เราสามารถคำนวณหาฟังก์ชันการแจกแจง (distribution function) ที่เกี่ยวข้องในทุกรูปแบบ ไม่ว่าจะเป็นการแจกแจงร่วม การแจกแจงตามขอบของทุกตัวแปร ในกรณีที่ทราบการแจกแจงย่อยเพียงสองอย่าง คือ การแจกแจงตามขอบและการแจกแจงแบบมีเงื่อนไข โดยมักจะใช้ร่วมกับกฎของความน่าจะเป็นรวม (Law of total probability)

ตัวอย่าง: การแจกแจงแบบมีเงื่อนไข

Example

สมมุติให้ค่า X คือค่าที่สุ่มจากการแจกแจงเอกรูป (uniform distribution) ในช่วง $[0,1]$ และ Y คือค่าที่สุ่มหลังจากที่ทราบว่าค่า $X=x$ จากการแจกแจงเอกรูป (uniform distribution) ในช่วง $[x,1]$ คำถามก็คือ ฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของ Y จะมีลักษณะอย่างไร?

- เนื่องจาก X มีการแจกแจงแบบเอกรูปในช่วง $[0,1]$ ดังนั้น เราสามารถเขียนฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของ X ได้เป็น

$$f_1(x) = \begin{cases} 1, & \text{สำหรับ } 0 \le x \le 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบมีเงื่อนไข (conditional p.d.f.) ของ Y เมื่อค่าของ $X=x$ เท่ากับ

$$g_2(y|x) = \begin{cases} \frac{1}{1-x}, & \text{สำหรับ } x \le y \le 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

วิธีทำ: การแจกแจงแบบมีเงื่อนไข

Example

- ใช้กฎการคูณเราจะพบว่า ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ของ (X, Y) เท่ากับ

$$f(x,y) = \begin{cases} \frac{1}{1-x}, & \text{สำหรับ } 0 < x \leq y \leq 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

- เราสามารถหาฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของ Y

$$f_2(y) = \begin{cases} \int_0^y \frac{1}{1-x} dx = -\ln(1-y) & \text{สำหรับ } 0 < y < 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

- เราสามารถหาฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบมีเงื่อนไข (conditional p.d.f.) ของ X เมื่อค่าของ $Y = y$ โดยประยุกต์ใช้ทฤษฎีบทของเบส์ (Bayes' theorem)

$$g_1(x|y) = \begin{cases} \frac{\frac{1}{1-x}}{-\ln(1-y)} = -\frac{1}{(1-x)\ln(1-y)} & \text{สำหรับ } 0 < x \leq y \leq 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

- ผลลัพธ์เดียวกันนี้สามารถหาได้จากการใช้นิยามของฟังก์ชันความหนาแน่นของความน่าจะเป็นแบบมีเงื่อนไข (conditional p.d.f.) เพราะเราทราบฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.)

ความสัมพันธ์ระหว่างความเป็นอิสระต่อกันและการแจกแจงแบบมีเงื่อนไข

Theorem

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่มีการแจกแจง $f_1(x)$ และ $f_2(y)$ ตามลำดับ ในทำนองเดียวกัน กำหนดให้ $g_1(x|y)$ และ $g_2(y|x)$ แทนการแจกแจงแบบมีเงื่อนไขของ X เมื่อค่าของ $Y = y$ และ การแจกแจงแบบมีเงื่อนไขของ Y เมื่อค่าของ $X = x$ ตามลำดับ ดังนั้น X และ Y เป็นอิสระต่อกัน (independent) ก็ต่อเมื่อ สำหรับทุกๆ ค่าของ y ที่ $f_2(y) > 0$ และทุกๆ ค่าของ x

$$g_1(x|y) = f_1(x) \tag{12}$$

และสำหรับทุกๆ ค่าของ x ที่ $f_1(x) > 0$ และทุกๆ ค่าของ y

$$g_2(y|x) = f_2(y) \tag{13}$$

ฟังก์ชันของตัวแปรสุ่ม (Functions of Random Variables)

- ตัวแปรพื้นฐานส่วนใหญ่มักจะอยู่ในรูปของ **ตัวแปรสุ่ม** เช่น รายได้, อัตราผลตอบแทน
- คำตอบที่ได้จากแบบจำลองของตัวแปรเหล่านี้เป็น **ตัวแปรสุ่ม** เช่นกัน
- ดังนั้น เพื่อให้เข้าใจคำตอบที่ได้ จึงจำเป็นจะต้องหาคุณสมบัติของคำตอบที่ได้ในฐานะที่เป็นตัวแปรสุ่ม
    - ต้องการหา “**การแจกแจงของฟังก์ชันของตัวแปรสุ่ม**”

ทฤษฎี: การแจกแจงของฟังก์ชันของตัวแปรสุ่มไม่ต่อเนื่อง

Theorem

กำหนดให้ X เป็นตัวแปรสุ่มไม่ต่อเนื่องที่มีฟังก์ชันความน่าจะเป็น (p.f.) f และ $Y = r(X)$ เป็น ***ฟังก์ชัน*** *ที่นิยามบนเซ็ตของค่าที่เป็นไปได้ทั้งหมดของ X ดังนั้น ฟังก์ชันความน่าจะเป็น (p.f.) ของ Y ซึ่งแทนด้วย g มีค่าเท่ากับ*

$$g(y) = Pr(Y = y) = Pr(r(X) = y) = \sum_{x:r(X)=y} f(x) \qquad (14)$$

สำหรับแต่ละค่า y ที่เป็นไปได้สำหรับตัวแปรสุ่ม Y

ทฤษฎี: ฟังก์ชันการแจกแจงของฟังก์ชันของตัวแปรสุ่มต่อเนื่อง

- การแจกแจงของตัวแปรสุ่มต่อเนื่องสามารถคำนวณได้โดยใช้หลักการสองขั้นตอน
    1. หาฟังก์ชันการแจกแจงสะสม (C.D.F.)
    2. หาฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) จากอนุพันธ์ของฟังก์ชันการแจกแจงสะสม (C.D.F.)

Theorem

กำหนดให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) f และ $Y = r(X)$ เป็นฟังก์ชันที่นิยามบนเซ็ตของค่าที่เป็นไปได้ทั้งหมดของ X ดังนั้น ฟังก์ชันการแจกแจงสะสม (C.D.F.) ของ Y ซึ่งแทนด้วย G มีค่าเท่ากับ

$$G(y) = Pr(Y \leq y) = Pr(r(X) \leq y) = \int_{\{x:r(X)\leq y\}} f(x)\,dx \tag{15}$$

สำหรับแต่ละค่า y ที่เป็นไปได้สำหรับตัวแปรสุ่ม Y และถ้าฟังก์ชันการแจกแจงสะสม (C.D.F.) ของ Y สามารถหาค่าอนุพันธ์ได้ ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) จะมีค่าเท่ากับ

$$g(y) = \frac{dG(y)}{dy} \tag{16}$$

ตัวอย่าง: การแจกแจงของฟังก์ชันของตัวแปรสุ่ม เมื่อ ตัวแปรสุ่ม X มีการแจกแจงเอกรูป

Example

สมมุติว่าตัวแปรสุ่ม X มีการแจกแจงเอกรูป (uniform distribution) ในช่วง $[0,1]$ แล้ว**ตัวแปรสุ่ม $Y = X^2$ มีการแจกแจงอย่างไร**?

- เนื่องจาก X มี**การแจกแจงแบบเอกรูป** ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ X มีค่าเท่ากับ

$$f(x) = \begin{cases} 1 \text{ สำหรับ } 0 \leq x \leq 1, \\ 0, \text{ สำหรับกรณีอื่น} \end{cases}$$

- ในช่วง $0 \leq x \leq 1$ แล้ว $0 \leq y \leq 1$ ดังนั้น เราจะต้องคำนวณหาฟังก์ชันการแจกแจงสะสม (C.D.F.) ของ Y สำหรับค่า $0 \leq y \leq 1$ ดังต่อไปนี้

$$\begin{aligned} G(y) = Pr(Y \leq y) = Pr\left(X^2 \leq y\right) &= Pr(-\sqrt{y} \leq X \leq \sqrt{y}) \\ = Pr(0 \leq X \leq \sqrt{y}) &= \int_0^{\sqrt{y}} f(x)\, dx = \sqrt{y}, \end{aligned}$$

  โดยหนึ่งในขั้นตอนการคำนวณได้ประยุกต์ใช้การที่ $f(x) = 0$ เมื่อ $x < 0$

- คำนวณหาค่าฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ได้จากอนุพันธ์ของ $G(y)$

$$g(y) = \begin{cases} \frac{1}{2\sqrt{y}} \text{ สำหรับ } 0 \leq y \leq 1, \\ 0, \text{ สำหรับกรณีอื่น} \end{cases}$$

ตัวอย่าง: การแจกแจงของฟังก์ชันของตัวแปรสุ่มเมื่อ ตัวแปรสุ่ม X มีการแจกแจงแบบปกติ

Example

สมมุติว่า X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ (normal distribution) ที่มีพารามิเตอร์ μ และ σ ซึ่งมีฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) เท่ากับ

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

คำถามก็คือ ตัวแปรสุ่ม $Y = aX + b$ ซึ่งเป็นฟังก์ชันเชิงเส้นของ X มีการแจกแจงอย่างไร?

- เช่นเดียวกับตัวอย่างก่อนหน้านี้ เราเริ่มจากการหาฟังก์ชันการแจกแจงสะสม (C.D.F.) ของ Y แล้วจึงหาฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) จากการหาอนุพันธ์
- ในทางเทคนิค เราจำเป็นต้องแยกออกเป็นสองกรณีคือ

วิธีทำ: การแจกแจงของฟังก์ชันของตัวแปรสุ่มเมื่อ ตัวแปรสุ่ม X มีการแจกแจงแบบปกติ กรณีที่ $a > 0$

Example

กรณีที่ $a > 0$:

- $$G(y) = Pr(Y \le y) = Pr(aX + b \le y) = Pr(aX + b \le y) = Pr\left(X \le \frac{y-b}{a}\right)$$
  $$= \int_{-\infty}^{\frac{y-b}{a}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$
- การหาอนุพันธ์
  $$g(y) = \frac{dG(y)}{dy} = \frac{d \int_{-\infty}^{\frac{y-b}{a}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}{dy} = \frac{1}{a\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}$$
  $$= \frac{1}{\sqrt{2\pi(a\sigma)^2}} e^{-\frac{(y-(a\mu+b))^2}{2(a\sigma)^2}}$$

วิธีทำ: การแจกแจงของฟังก์ชันของตัวแปรสุ่มเมื่อ ตัวแปรสุ่ม X มีการแจกแจงแบบปกติ กรณีที่ $a < 0$

Example

กรณีที่ $a < 0$:

- $$G(y) = Pr(Y \le y) = Pr(aX + b \le y) = Pr(aX + b \le y) = Pr\left(X \ge \frac{y-b}{a}\right)$$
  $$= 1 - \int_{-\infty}^{\frac{y-b}{a}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$
- การหาอนุพันธ์
  $$g(y) = \frac{dG(y)}{dy} = -\frac{d\int_{-\infty}^{\frac{y-b}{a}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx}{dy} = \frac{1}{a\sqrt{2\pi\sigma^2}} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}}$$
  $$= \frac{1}{\sqrt{2\pi(a\sigma)^2}} e^{-\frac{(y-(a\mu+b))^2}{2(a\sigma)^2}}$$

ฟังก์ชันของตัวแปรสุ่มที่มีการแจกแจงแบบปกติ (normal distribution)

- ตัวแปรสุ่ม $Y = aX + b$ ซึ่งเป็นฟังก์ชันเชิงเส้นของ X มีการแจกแจงแบบปกติ (normal distribution) ที่มีพารามิเตอร์ $a\mu + b$ และ $a\sigma$
- **บทเรียน**: ผลลัพธ์ที่ได้จากกรณีที่ $a > 0$ และ $a < 0$ นั้นเหมือนกัน แสดงว่า ค่าของพารามิเตอร์ a ที่มีผลต่อการแปลงการแจกแจงในกรณีที่เป็นฟังก์ชันเชิงเส้นคือ "**ค่าสมบูรณ์ (absolute value) ของ** a"

ทฤษฎี: เมื่อฟังก์ชันของตัวแปรสุ่มเป็นฟังก์ชันเชิงเส้น

Theorem

สมมุติให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มี f เป็นฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) และตัวแปรสุ่ม $Y = aX + b$ โดยที่ $a \neq 0$ แล้ว ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ Y เท่ากับ

$$g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right), \text{ สำหรับ } -\infty < y < \infty \tag{17}$$

- ทฤษฎีบทนี้เป็นกรณีพิเศษของทฤษฎีบทที่กล่าวถึงการแปลงการแจกแจงด้วยฟังก์ชันหนึ่งต่อหนึ่งที่หาอนุพันธ์ได้ (one-to-one and differentiable function)

ทฤษฎี: การแปลงค่าของฟังก์ชันของตัวแปรสุ่ม

Theorem

กำหนดให้ X เป็นตัวแปรสุ่มต่อเนื่องที่มี f เป็นฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) โดยที่ $\Pr(a < X < b) = 1$ และตัวแปรสุ่ม $Y = r(X)$ โดยที่ฟังก์ชัน $r(x)$ เป็นฟังก์ชันหนึ่งต่อหนึ่งที่หาอนุพันธ์ได้ (one-to-one and differentiable function) สำหรับ $a < x < b$ ซึ่งมีผลทำให้เกิดภาพฉาย[a] (image) ของช่วง (a, b) เท่ากับ (α, β)
กำหนดให้ $s(y)$ คือฟังก์ชันผกผัน (inverse function) ของ $r(x)$ นั่นคือ $x = s(y)$ ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ Y มีค่าเท่ากับ

$$g(y) = \begin{cases} f(r(x)) \left| \frac{ds(y)}{dy} \right| & \text{สำหรับ } \alpha \leq y \leq \beta, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{18}$$

[a] หากพิจารณาอย่างไม่เป็นทางการ ภาพฉาย (image) หมายถึงเซ็ตของจำนวนจริงที่เป็นผลลัพธ์ของฟังก์ชัน r ที่เป็นไปได้ทั้งหมด หรือในทางเทคนิค เราอาจจะเขียนได้เป็น $r : (a, b) \to (\alpha, \beta)$

การพิสูจน์: การแปลงค่าของฟังก์ชันของตัวแปรสุ่ม

Proof.

การพิสูจน์แบ่งออกเป็นสองส่วน ดังต่อไปนี้

1. กรณีที่ r เป็นฟังก์ชั่นที่เพิ่มขึ้น (increasing function): ผลที่ตามมาก็คือ s ก็เป็นฟังก์ชั่นที่เพิ่มขึ้น ซึ่งช่วยให้เราเขียนฟังก์ชันการแจกแจงสะสม (C.D.F.) ของ Y สำหรับ $y \in (\alpha, \beta)$

$$G(y) = Pr(Y \leq y) = Pr(r(X) \leq y) = Pr(X \leq s(y)) = F(s(y))$$

เราสามารถประยุกต์ใช้กฎลูกโซ่ (chain rule) เพื่อหาอนุพันธ์ของ G ได้ดังต่อไปนี้

$$g(y) = \frac{dG(y)}{dy} = \frac{dF(s(y))}{dy} = \frac{dF(s(y))}{dx}\frac{ds(y)}{dy} = f(s(y))\frac{ds(y)}{dy} = f(s(y))\left|\frac{ds(y)}{dy}\right|$$

2. กรณีที่ r เป็นฟังก์ชั่นที่ลดลง (decreasing function): ผลที่ตามมาก็คือ s ก็เป็นฟังก์ชั่นที่ลดลง ซึ่งช่วยให้เราเขียนฟังก์ชันการแจกแจงสะสม (C.D.F.) ของ Y สำหรับ $y \in (\alpha, \beta)$

$$G(y) = Pr(Y \leq y) = Pr(r(X) \leq y) = Pr(X \geq s(y)) = 1 - F(s(y))$$

เราสามารถประยุกต์ใช้กฎลูกโซ่ (chain rule) เพื่อหาอนุพันธ์ของ G ได้ดังต่อไปนี้

$$g(y) = \frac{dG(y)}{dy} = -\frac{dF(s(y))}{dy} = -\frac{dF(s(y))}{dx}\frac{ds(y)}{dy} = -f(s(y))\frac{ds(y)}{dy} = f(s(y))\left|\frac{ds(y)}{dy}\right|$$

□

ตัวอย่าง: การแปลงค่าของฟังก์ชันของตัวแปรสุ่ม

Example

สมมุติให้ X เป็นตัวแปรสุ่มที่บอกถึงอัตราการให้บริการลูกค้าที่เข้าแถวรอของธนาคารแห่งหนึ่ง โดยกำหนดให้ f แทนฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ X
คำถามที่สนใจก็คือ การแจกแจงของเวลาที่ใช้ในการรอ (waiting time) เป็นอย่างไร?

- กำหนดให้ เวลาที่ใช้ในการรอ (waiting time) Y เป็นส่วนกลับของอัตราการให้บริการ นั่นคือ $Y = \frac{1}{X}$ โดยนิยามฟังก์ชัน $r(x) = \frac{1}{x}$ ดังนั้น $s(y) = \frac{1}{y}$ และอนุพันธ์ของ s เท่ากับ

$$\frac{ds(y)}{dy} = -\frac{1}{y^2}$$

- ซึ่งมีค่าน้อยกว่าศูนย์และ $\left|\frac{ds(y)}{dy}\right| = \frac{1}{y^2}$ ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ Y เท่ากับ

$$g(y) = f\left(\frac{1}{y}\right)\frac{1}{y^2}$$

ฟังก์ชันของตัวแปรสุ่มสองตัวขึ้นไป (Functions of Two or More Random Variables)

- หลักการที่ใช้ในการแปลงการแจกแจงในกรณีของฟังก์ชันของตัวแปรสุ่มสองตัวขึ้นไปก็ยังเป็นเช่นเดิม
  1. หาฟังก์ชันการแจกแจงสะสม (C.D.F.)
  2. หาฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) จากอนุพันธ์ของฟังก์ชันการแจกแจงสะสม (C.D.F.)
  3. มีความแตกต่างกันบ้างในทางเทคนิค โดยเฉพาะสูตรที่ใช้ในการแปลงโดยตรง (ในทำนองเดียวกับทฤษฎีบทที่แล้ว) ที่จำเป็นต้องใช้เครื่องมือที่เรียกว่าจาโคเบียนเมตริกซ์ (Jacobian matrix)

ทฤษฎี: การแจกแจงของฟังก์ชันของตัวแปรสุ่มสองตัวขึ้นไป (Functions of Two or More Random Variables)

Theorem

กำหนดให้ $X_1, \ldots, X_n$ *เป็นตัวแปรสุ่มต่อเนื่องที่มี* f *แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.)*
สมมุติให้เซ็ต $S \subset \mathbb{R}^n$ *โดยที่* $Pr\left((X_1, \ldots, X_n) \in S\right) = 1$
พิจารณาตัวแปรสุ่ม $Y_1, \ldots, Y_n$ *ซึ่งเป็นฟังก์ชันของ* $X_1, \ldots, X_n$ *ดังนี้*

$$\begin{aligned} Y_1 &= r_1\left(X_1, \ldots, X_n\right) \\ &\vdots \\ Y_n &= r_n\left(X_1, \ldots, X_n\right) \end{aligned} \tag{19}$$

โดยที่ r_i *สำหรับ* $i = 1, \ldots, n$ *เป็นฟังก์ชันหนึ่งต่อหนึ่งที่หาอนุพันธ์ได้ (one-to-one and differentiable function) ซึ่งแปลงค่าจากเซ็ต* S *ไปสู่เซ็ต* $T \subset \mathbb{R}^n$ *และกำหนดให้ฟังก์ชันผกผัน* s_i *ที่สอดคล้องกับฟังก์ชัน* r_i *สำหรับ* $i = 1, \ldots, n$ *เท่ากับ*

$$\begin{aligned} x_1 &= s_1\left(y_1, \ldots, y_n\right) \\ &\vdots \\ x_n &= s_n\left(y_1, \ldots, y_n\right) \end{aligned} \tag{20}$$

ทฤษฎี: การแจกแจงของฟังก์ชันของตัวแปรสุ่มสองตัวขึ้นไป (Functions of Two or More Random Variables) con't

Theorem

ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ของ $Y_1, \ldots, Y_n$ *หาได้จาก*

$$g(y_1, \ldots, y_n) = \begin{cases} f(s_1, \ldots, s_n)\,|J| & \text{สำหรับ } (y_1, \ldots, y_n) \in T, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{21}$$

โดยที่ J *แทนดีเทอร์มิแนนต์ (determinant) ของจาโคเบียนเมตริกซ์ (Jacobian matrix) ซึ่งมีค่าเท่ากับ*

$$J = \det \begin{bmatrix} \frac{\partial s_1}{\partial y_1} & \cdots & \frac{\partial s_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial s_n}{\partial y_1} & \cdots & \frac{\partial s_n}{\partial y_n} \end{bmatrix} \tag{22}$$

ในที่นี้ สัญลักษณ์ det *หมายถึงดีเทอร์มิแนนต์ (determinant) ของเมตริกซ์จตุรัส*

ตัวอย่าง: การแจกแจงของฟังก์ชันของตัวแปรสุ่มสองตัวขึ้นไป

Example

พิจารณาตัวแปรสุ่มต่อเนื่อง X และ Y ซึ่งมีฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) เท่ากับ

$$f(x_1, x_2) = \begin{cases} 4xy \text{ สำหรับ } 0 < x_1 < 1 \text{ และ } 0 < y < 1, \\ 0, \text{ สำหรับกรณีอื่น} \end{cases}$$

สิ่งที่ต้องการทราบในตัวอย่างนี้คือ ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ตัวแปรสุ่ม $Z_1 = \frac{X}{Y}$ และ $Z_2 = XY$

- เราสามารถเขียนฟังก์ชันที่เราสนใจได้เป็น

$$z_1(x, y) = \frac{x}{y} \text{ และ } z_2(x, y) = xy$$

- เราสามารถหาฟังก์ชันผกผันของฟังก์ชันทั้งสอง

$$x_1 = s_1(z_1, z_2) = \sqrt{\frac{z_1}{z_2}} \text{ และ } x_2 = s_2(z_1, z_2) = \sqrt{z_1 z_2}$$

Example

- S ในตัวอย่างนี้เท่ากับ $S = \left\{(x,y) \in \mathbb{R}^2 : 0 < x_1 < 1 \text{ และ } 0 < y < 1\right\}$ เราหาเซ็ต T ได้ดังต่อไปนี้ ก่อนอื่น เนื่องจาก $x > 0, y > 0, z_1 = \frac{x}{y}$ และ $z_2 = xy$ ดังนั้น $z_1, z_2 > 0$
- เราสามารถประยุกต์ใช้ฟังก์ชันผกผันร่วมกับเงื่อนไขของ x และ y ที่กำหนดให้ $0 < x_1 < 1$ และ $0 < y < 1$ ในการคำนวณหาเซ็ต $T = \left\{(z_1, z_2) \in \mathbb{R}^2 : 0 < z_2 < \frac{1}{z_1} \text{ และ } 0 < z_2 < z_1\right\}$
- หาดีเทอร์มิแนนต์ (determinant) ของจาโคเบียนเมตริกซ์ (Jacobian matrix)

$$J = \det \begin{bmatrix} \frac{\partial s_1}{\partial z_1} & \frac{\partial s_1}{\partial z_2} \\ \frac{\partial s_2}{\partial z_1} & \frac{\partial s_2}{\partial z_2} \end{bmatrix} = \det \begin{bmatrix} \frac{1}{2}\sqrt{\frac{z_2}{z_1}} & \frac{1}{2}\sqrt{\frac{z_1}{z_2}} \\ -\frac{1}{2}\sqrt{\frac{z_2}{z_1^3}} & \frac{1}{2}\sqrt{\frac{1}{z_1 z_2}} \end{bmatrix} = \frac{1}{2z_1}$$

- เนื่องจาก $z_1 > 0$ ดังนั้น $|J| = \frac{1}{2z_1}$ และฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ตัวแปรสุ่ม $Z_1 = \frac{X}{Y}$ และ $Z_2 = XY$ เท่ากับ

$$g(z_1, z_2) = \begin{cases} 2\frac{z_2}{z_1} & \text{สำหรับ } (z_1, z_2) \in T, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

ทฤษฎี: การแจกแจงของฟังก์ชันของตัวแปรสุ่มสองตัวขึ้นไปของฟังก์ชันเชิงเส้น

Theorem

กำหนดให้ $\mathbf{X} = (X_1, \ldots, X_n)$ *เป็นตัวแปรสุ่มต่อเนื่องที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.)* f *และนิยามฟังก์ชันเชิงเส้น*

$$Y = AX \tag{23}$$

โดยที่ A *คือเมทริกซ์ไม่เอกฐาน (nonsingular matrix) ที่มีขนาด* $n \times n$ *(บางครั้งเราเรียกค่าจำนวนจริงในเมทริกซ์นี้ว่าค่าสัมประสิทธิ์) ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ของ* Y *หาได้จาก*

$$g(y) = \frac{1}{|detA|} f\left(A^{-1}y\right) \tag{24}$$

โดยที่ A^{-1} *แทนส่วนผกผันของ* A

สรุป: การหาการแจกแจงของฟังก์ชันของตัวแปรสุ่มสองตัวขึ้นไป

- ทฤษฎีบทสองอันหลังนั้นมีประโยชน์ต่อการหาการแจกแจงของฟังก์ชันของตัวแปรสุ่มอย่างมาก
- มีข้อจำกัดในบางประการ
  - จำนวนฟังก์ชันที่พิจารณาต้องเท่ากับจำนวนตัวแปรสุ่มพื้นฐาน เพื่อให้จาโคเบียนเมตริกซ์ (Jacobian matrix) เป็นจัตุรัสจึงจะหาค่าได้
  - เราอาจจะสนใจปัญหาที่มีจำนวนฟังก์ชันแตกต่างจากจำนวนตัวแปรสุ่มพื้นฐาน ซึ่งโดยมากมักจะมีจำนวนฟังก์ชันน้อยกว่าจำนวนตัวแปร
- ยกตัวอย่างเช่น การประมาณการณ์สมการเชิงเส้นอย่างง่ายมักจะมีฟังก์ชันเดียวแต่มีตัวแปรสุ่มพื้นฐานจำนวนมาก
  - กรณีเช่นนี้เราสามารถกลับไปใช้หลักพื้นฐานที่เริ่มจากการหาฟังก์ชันการแจกแจงสะสม (C.D.F.)
  - แล้วจึงหาฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) จากอนุพันธ์ของฟังก์ชันการแจกแจงสะสม (C.D.F.)

ตัวอย่าง: การหาการแจกแจงของฟังก์ชันของตัวแปรสุ่มหลายตัวในกรณีที่ไม่สามารถหาจาโคเบียนเมตริกซ์ (Jacobian matrix) ได้

Example

พิจารณาตัวอย่างสุ่ม (random sample) ขนาด n ซึ่งสุ่มเลือกมาจากตัวแปรสุ่ม $X_1, X_2, \ldots, X_n$ โดยที่แต่ละตัวมีฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) f และฟังก์ชันการแจกแจงสะสมร่วม (C.D.F.) F ที่เหมือนกัน พิจารณาตัวแปรสุ่มที่แทนค่าสูงสุด Y_m และค่าต่ำสุด Y_0 ของตัวแปรสุ่มเหล่านั้นดังต่อไปนี้

$$
\begin{aligned}
Y_m &= \max\left\{X_1, X_2, \ldots, X_n\right\}, \\
Y_0 &= \min\left\{X_1, X_2, \ldots, X_n\right\}
\end{aligned}
$$

สิ่งที่ต้องการทราบคือ การแจกแจงของ Y_m และ Y_0 ซึ่งสามารถคำนวณได้ตามขั้นตอนต่อไปนี้

- กำหนดฟังก์ชันการแจกแจงสะสม (C.D.F.) ของค่าสูงสุด Y_m

$$G_m(y) = Pr(Y_m \le y) = Pr(X_1 \le y, X_2 \le y, \ldots, X_n \le y)$$

- การที่ข้อมูลเป็นแบบตัวอย่างสุ่ม (random sample) ทำให้ทราบว่า $X_1, X_2, \ldots, X_n$ เป็นอิสระต่อกัน (independent) ดังนั้น

$$
\begin{aligned}
G_m(y) &= Pr(X_1 \le y, X_2 \le y, \ldots, X_n \le y) = Pr(X_1 \le y)\, Pr(X_2 \le y) \cdots Pr(X_n \le y) \\
&= [F(y)]^n
\end{aligned}
$$

วิธีทำ: การหาการแจกแจงของฟังก์ชันของตัวแปรสุ่มหลายตัวในกรณีที่ไม่สามารถหาจาโคเบียนเมตริกซ์ (Jacobian matrix) ได้

Example

- ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ Y_n หาได้จากอนุพันธ์ของ $G_n(y)$ ดังนี้

$$g_m(y) = \frac{dG_m(y)}{dy} = nf(y)\left[F(y)\right]^{n-1}$$

- การแจกแจงของฟังก์ชันค่าต่ำสุด Y_0 เป็นดังนี้

$$\begin{aligned}
G_0(y) &= Pr(Y_0 \leq y) = 1 - Pr(Y_0 > y) = 1 - Pr(X_1 > y, X_2 > y, \ldots, X_n > y) \\
&= 1 - Pr(X_1 > y)\, Pr(X_2 > y) \cdots Pr(X_n > y) \\
&= 1 - \left[1 - F(y)\right]^n \\
g(y) &= \frac{dG(y)}{dy} = nf(y)\left[1 - F(y)\right]^{n-1}
\end{aligned}$$

- การแจกแจงร่วม (joint distribution) ของ (Y_0, Y_m) ได้ดังต่อไปนี้

$$G(y_0, y_m) = Pr(Y_0 \leq y_0, Y_m \leq y_n)$$

วิธีทำ: การหาการแจกแจงของฟังก์ชันของตัวแปรสุ่มหลายตัวในกรณีที่ไม่สามารถหาจาโคเบียนเมตริกซ์ (Jacobian matrix) ได้

Example

- กำหนดให้ A แทนเหตุการณ์ที่ $Y_m \leq y_m$ นั่นคือเซ็ต $\{(x_1, \ldots, x_n) : Y_m \leq y_m\}$ และ B แทนเหตุการณ์ที่ $Y_0 \leq y_0$
- นั่นคือเซ็ต $\{(x_1, \ldots, x_n) : Y_0 \leq y_0\}$ และจากทฤษฎีบทที่ผ่านๆ มากล่าวว่า $Pr(A \cap B) = Pr(A) - Pr(A \cap B^c)$ เราสามารถเขียน G ได้เป็น

$$\begin{aligned} G(y_0, y_m) &= Pr(Y_0 \leq y_0, Y_m \leq y_m) = Pr(Y_m \leq y_m) - Pr(Y_m \leq y_m, Y_0 > y_0) \\ &= G_m(y_m) - Pr(y_0 < X_1 \leq y_m, \ldots, y_0 < X_n \leq y_m) \\ &= G_m(y_m) - \Pi_{i=1}^{n} Pr(y_0 < X_i \leq y_m) = [F(y_m)]^n - [F(y_m) - F(y_0)]^n \end{aligned}$$

และฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ในกรณีที่ $-\infty < y_0 < y_m < \infty$ สามารถหาได้จากอนุพันธ์ของ $G(y_0, y_m)$ ดังนี้

$$g(y_0, y_m) = \frac{\partial^2 G(y_0, y_m)}{\partial y_0 \partial y_m} = n(n-1) f(y_0) f(y_m) [F(y_m) - F(y_0)]^{n-2} \tag{25}$$

ส่วนในกรณีอื่นๆ จะกำหนดให้ $g(y_0, y_m) = 0$

ตัวอย่าง: การแจกแจงของค่าพิสัย (range)

Example

พิจารณาฟังก์ชัน $Z = Y_m - Y_0$ ซึ่งเรียกว่าค่าพิสัย (range) คำถามก็คือ การแจกแจงของค่าพิสัยเป็นอย่างไร?

- เราทราบแล้วว่าฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ของตัวแปรสุ่ม (Y_0, Y_m) เท่ากับ

$$g(y_0, y_m) = \frac{\partial^2 G(y_0, y_m)}{\partial y_0 \partial y_m} = n(n-1) f(y_0) f(y_m) \left[F(y_m) - F(y_0)\right]^{n-2}$$

- เราสามารถเทียบเคียงฟังก์ชันค่าพิสัย Z กับฟังก์ชันเชิงเส้นในตัวอย่างที่ 3.31 (ใน Lecture Note) ได้โดยกำหนดให้ $a_1 = -1, a_2 = 1,$ และ $b = 0$ ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ Z เท่ากับ

$$h(z) = \int_{-\infty}^{\infty} g(y_n - z, y_n)\, dy_n$$

- เรายังสามารถประยุกต์ใช้หลักการเปลี่ยนตัวแปรการอินทิเกรต (change of variable) โดยกำหนดให้ $w = y_n - z$ ดังนั้น $dy_n = dw$ และ $y_n = w + z$ เพื่อแปลงฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ Z เป็น

$$h(z) = \int_{-\infty}^{\infty} g(w, w+z)\, dw$$