บทที่ 1
ทฤษฎีความน่าจะเป็น

การทดลองทางสถิติหรือค่าสังเกตในตัวอย่างซึ่งเราต้องการข้อสรุปเกี่ยวกับประชากรที่เราเลือกตัวอย่างมา ก่อนเราสามารถทำจริง ๆ เราต้องพัฒนาตัวแบบคณิตศาสตร์ของประชากรนี้จะต้องการทฤษฎีของความน่าจะเป็นทางคณิตศาสตร์ ดังนั้น จึงเห็นได้ว่า ทฤษฎีนี้เป็นพื้นฐานสำคัญต่อสถิติคณิตศาสตร์

ข้อต่อไปนี้ เราควรจำว่าตัวแบบคณิตศาสตร์ซึ่งช่วยในการแก้ปัญหาชีวิตจริง ที่ใช้ในสาขาต่าง ๆ ตัวอย่างเช่น ในวิชาว่าด้วยกำลังและความเคลื่อนไหวของอากาศ เราอาจเริ่มจากกฎพื้นฐานของวิชากลศาสตร์ และหาได้จากตัวแบบคณิตศาสตร์ของการเคลื่อนที่ของเครื่องบิน ตัวแบบคณิตศาสตร์ควรเป็นชนิดง่ายและเป็นไปได้

เรามาพิจารณาทฤษฎีความน่าจะเป็นพื้นฐาน โดยเฉพาะแนวทางการศึกษาของความน่าจะเป็นคณิตศาสตร์แล้ว เราจะสืบถึงการแจกแจงความน่าจะเป็นที่สำคัญที่สุด สิ่งเหล่านี้เป็นการแจกแจงทางทฤษฎีที่ใช้เป็นตัวแบบของประชากรในการเปรียบเทียบกับการแจกแจงตัวอย่างโดยอาศัยความสังเกต

1.1 แนวทางการศึกษาพื้นฐาน

ทฤษฎีความน่าจะเป็นทางคณิตศาสตร์ ได้กำเนิดมาจากทฤษฎีของโอกาสหรือหนทาง Pascal และ Fermat ได้ทำการศึกษาความน่าจะเป็นของปัญหา ต่อมา Jakob Bernoulli, Abraham de Moivre และ Pierre Simon de Laplace ได้พัฒนาทฤษฎีของระบบ (การแจกแจงต่าง ๆ) ทฤษฎีความน่าจะเป็นมีความสำคัญในวิชาดาราศาสตร์และวิชาคำนวณที่ว่าด้วยรูปร่างและเนื้อที่ของโลกในการเกี่ยวข้องกับทฤษฎีของความคลาดเคลื่อนจากผลที่วัดได้ Gauss และ Laplace ได้สร้างการแจกแจงที่สำคัญนี้ สำหรับรายละเอียดเพิ่มเติมจนถึงกึ่งศตวรรษที่สิบเก้าก็ให้นักศึกษาไปศึกษาได้ใน Todhunter ค.ศ. 1949 (cf Appendix 3)

1.2 การทดลองเชิงสุ่มและผลลัพธ์

ในหัวข้อนี้ เราจะบรรยายหลักพื้นฐานง่าย ๆ ซึ่งก็กล่าวมาก่อนบ้างแล้ว

การทดลองเชิงสุ่มหรือการสังเกตโดยสุ่ม ในที่นี้เราหมายถึงกระบวนการที่มีคุณสมบัติดังต่อไปนี้

1. การดำเนินการจะต้องเป็นไปตามกฎของเซทที่กำหนดไว้
2. กระบวนการกระทำกันซ้ำ ๆ กันโดยธรรมชาติหรือสามารถคิดขึ้น
3. ผลลัพธ์ของแต่ละการดำเนินการขึ้นอยู่กับโอกาส (นั่นคือ เกี่ยวกับอิทธิพลซึ่งเราไม่สามารถควบคุมได้) ไม่สามารถทำนายได้อย่างเดียวเท่านั้น

ผลลัพธ์ของการดำเนินการครั้งเดียวของการทดลอง เรียกว่า outcome ของการทดลอง

ตัวอย่างคือเกมต่าง ๆ อย่างเช่น การทอดลูกเต๋าหนึ่งลูก โยนเหรียญหนึ่งอัน หรือการหยิบไพ่สองใบจากไพ่สำรับหนึ่ง เทคนิคของการทดลอง อย่างเช่น การเลือกและตรวจสอบสกรู 10 ตัว โดยสุ่มจากกล่องใบหนึ่งมีสกรู 100 ตัว การวัดอายุของหลอดไฟโดยสุ่มเลือกมา 50 หลอดจากทั้งหมด การคำนวณผลของกระบวนการทางเคมีภายใต้เงื่อนไขต่าง ๆ การทดลองเกี่ยวกับการเลี้ยงหมูหลาย ๆ ตัว ในอัตราต่าง ๆ กัน โดยการดำเนินการครั้งเดียวกับสัตว์ชนิดเดียวกัน การสุ่มเลือกบุคคล 20 คน จากชนกลุ่มหนึ่ง และวัดความดันของชายเหล่านั้น หรือความเห็นของเขาเกี่ยวกับภาพยนตร์ ฯลฯ หรือการทดลองอื่น ๆ ที่มีคุณสมบัติข้างต้น

1.3 เซททางพีชคณิต

ก่อนที่จะกล่าวถึงแนวทางของตัวแบบความน่าจะเป็น ซึ่งเราประสงค์ที่จะสร้างหรือพัฒนา จึงจำเป็นที่จะต้องมีแนวทางหรือความคิดเกี่ยวกับทฤษฎีทางคณิตศาสตร์ของเซท ซึ่งมีความสำคัญมาก อย่างไรก็ตาม เราต้องการแนวทางพื้นฐานเพียงเล็กน้อยเท่านั้น

เซท เซทเป็นที่รวบรวมของวัตถุหรือสิ่งของ เซทโดยทั่ว ๆ ไปใช้อักษรตัวใหญ่ A, B... ฯลฯ มีอยู่สามวิธีด้วยกันที่ใช้บรรยายว่า วัตถุหรือสิ่งของชนิดไหนที่อยู่ในเซท

ก. เราอาจเขียนจำนวนหมายเลขของ A อย่างเช่น $A = \{1, 2, 3, 4\}$ บรรยายถึงเซทว่าประกอบด้วยเลขจำนวนเต็มบวก 1, 2, 3 และ 4

ข. เราอาจบรรยายเซท A ในรูปของคำ อย่างเช่น เราอาจกล่าวว่า A ประกอบด้วยเลขจำนวนจริงทั้งหมด ตั้งแต่ 0 ถึง 1

ค. เราอาจบรรยายเซทข้างต้น (ข้อ ข.) ด้วยการเขียน $A = \{x \mid 0 \le x \le 1\}$ นั่นคือ A เป็นเซทของค่า x ทั้งหมด ในเมื่อ x เป็นเลขจำนวนจริงตั้งแต่ 0 ถึง 1

วัตถุแต่ละชิ้นที่ประกอบขึ้นเป็นเซท A เรียกว่า สมาชิกของ A เมื่อไร "a" เป็นสมาชิกของ A เราเขียนได้ $a \in A$ และเมื่อไร "a" ไม่เป็นสมาชิกของ A เราเขียนได้ $a \notin A$ มีเซทเฉพาะอยู่สองเซท ที่สนใจคือ เซทที่มีสมาชิกจำนวนจำกัดของวัตถุ กับเซทที่มีสมาชิกจำนวนไม่จำกัด ตัวอย่างเช่น เราอาจเกี่ยวข้องกับเลขจำนวนจริงทั้งหมด จำนวนวัตถุทั้งหมดที่ได้จากการผลิตระยะ 24 ชั่วโมง ฯลฯ เรานิยามเซทจักรวาล (Universal set) ได้เป็นเสมือนเซทของวัตถุทั้งหมด ภายใต้การพิจารณา เซทนี้โดยทั่ว ๆ ไปเขียนได้เป็น U

ส่วนเซทอื่น ๆ ที่ไม่มีค่าเลย อย่างเช่น เซท A อาจบรรยายได้เสมือนเซทของเลขจำนวนจริงทั้งหมด x ที่สอดคล้องสมการ $x^2 + 1 = 0$ เราทราบว่าไม่มีสมาชิกเช่นนั้น นั่นคือเซท A ไม่มีสมาชิกเลย สภาวะนี้เกิดขึ้นบ่อย ๆ และจัดเข้าในรูปของเซทที่มีชื่อเฉพาะ ด้วยเหตุนี้เราจึงนิยามเซทว่างเปล่าเป็นเซทที่ไม่มีสมาชิกเลย ใช้สัญลักษณ์ $\varnothing$

เมื่อไรเราพิจารณาสองเซท A กับ B โดยที่สมาชิกของ A เป็นสมาชิกของ B แล้วเราเรียกว่า A เป็นเซทย่อยของ B เราเขียนได้ $A \subset B$ ในความหมายเดียวกันก็อาจกำหนดได้ต่อ $B \subset A$ เซททั้งสองเหมือนกัน $A = B$ ก็ต่อเมื่อ $A \subset B$ และ $B \subset A$ นั่นคือ สองเซทเท่ากันก็ต่อเมื่อมีสมาชิกเหมือนกัน

คุณสมบัติของเซทว่างเปล่ากับเซทจักรวาล มีดังนี้

ก. ทุก ๆ เซท A เราได้ $\varnothing \subset A$

ข. ทุก ๆ เซท A ที่อยู่ในเซทจักรวาล U แล้ว เราได้ $A \subset U$

ตัวอย่างที่ 1 สมมติว่า U = เลขจำนวนจริงทั้งหมด $A = \{x \mid x^2 + 2x - 3 = 0\}$, $B = \{x \mid (x - 2)(x^2 + 2x - 3) = 0\}$ และ $C = \{x \mid x = -3, 1, 2\}$ แล้ว $A \subset B$ และ $B = C$

ตัวอย่างที่ 2 ให้ $A_1 = \{(x, y) ; 0 \le x = y \le 1\}$ **และ** $A_2 = \{(x, y) ; 0 \le x \le 1, 0 \le y \le 1\}$ เนื่องจากว่าสมาชิกของ A_1 เป็นจุดบนเส้นทะแยงของสี่เหลี่ยมจัตุรัส ดังนั้น $A_1 \subset A_2$

นิยาม เซทของสมาชิกทั้งหมดซึ่งเป็นสมาชิกอย่างน้อย หนึ่งของเซท A_1 กับ A_2 เรียกว่าการรวม (Union) ของ A_1 กับ A_2 การรวมของ A_1 กับ A_2 แสดงได้เป็น $A_1 \cup A_2$ การรวมของหลาย ๆ เซท $A_1, A_2, A_3, \ldots$ เป็นเซทของสมาชิกทั้งหมด ที่เป็นอย่างน้อย หนึ่งของหลาย ๆ เซท

การรวมนี้แสดงได้โดย $A_1 \cup A_2 \cup A_3 \cup \ldots$ หรือ $A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_k$ ถ้าหากว่า k เป็นเลขจำนวนจำกัดของเซทที่เกี่ยวข้อง

ตัวอย่างที่ 2 ให้ $A_1 = \{x ; x = 0, 1, \ldots 10\}$ และ $A_2 = \{x ; x = 8, 9, 10, 11$ หรือ $11 < x \leq 12\}$ แล้ว $A_1 \cup A_2 = \{x; x = 0, 1, \ldots, 8, 9, 10, 11$ หรือ $11 < x \leq 12\}$

$$= \{x ; x = 0, 1, \ldots, 8, 9, 10, 11 \leq X \leq 12\}$$

ตัวอย่างที่ 3 ให้ $A_2 = \varnothing$ แล้ว $A_1 \cup A_2 = A_1$ สำหรับทุก ๆ เซท A_1

ตัวอย่างที่ 4 สำหรับทุก ๆ เซท A $A \cup A = A$

ตัวอย่างที่ 5 ให้ $A_k = \{x ; 1/(k+1) \leq x \leq 1\}, k = 1, 2, 3, \ldots$ แล้ว $A_1 \cup A_2 \cup A_3 \cup \ldots = \{x ; 0 < x \leq 1\}$ สังเกตว่าเลขศูนย์ไม่ได้อยู่ในเซท เนื่องจากว่าไม่ได้อยู่ในหนึ่งของหลาย ๆ เซท $A_1, A_2, \ldots$

นิยาม เซทของสมาชิกทั้งหมดซึ่งเป็นสมาชิกของแต่ละเซท A_1 กับ A_2 เรียกว่า การร่วม (Intersection) ของ A_1 และ A_2 การร่วมของ A_1 และ A_2 แสดงได้โดย $A_1 \cap A_2$ การร่วมของหลายๆ เซท $A_1, A_2, A_3, \ldots$ เป็นเซทของสมาชิกทั้งหมดที่เป็นสมาชิกของแต่ละเซท $A_1, A_2, \ldots$ การร่วมนี้แสดงได้โดย $A_1 \cap A_2 \cap A_3 \cap \ldots$ หรือ $A_1 \cap A_2 \cap \ldots \cap A_k$ ถ้าหากว่า k เป็นหมายเลขของเซทที่เกี่ยวข้อง

ตัวอย่างที่ 6 ให้ $A_1 = \{(x, y) ; (x, y) = (0, 0), (0, 1), (1, 1)\}$ และ $A_2 = \{(x, y) ; (x, y) = (1, 1), (1, 2), (2, 1)\}$ ดังนั้น

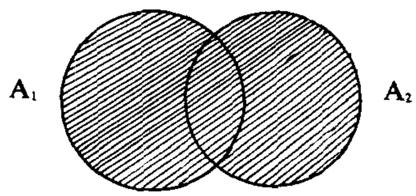
$$A_1 \cap A_2 = \{(x, y) ; (x, y) = (1, 1)\}$$

ตัวอย่างที่ 7 ให้ $A_1 = \{(x, y) ; 0 \leq x+y \leq 1\}$ และ $A_2 = \{(x, y) ; 1 < x+y\}$ ดังนั้น A_1 และ A_2 ไม่มีจุดร่วม และ $A_1 \cap A_2 = \varnothing$

ตัวอย่างที่ 8 สำหรับทุก ๆ เซท A, $A \cap A = A$ กับ $A \cap \varnothing = \varnothing$

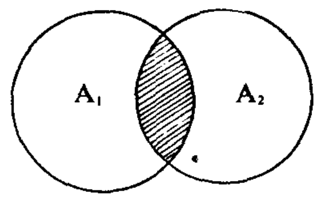
ตัวอย่างที่ 9 ให้ $A_k = \{x ; 0 < x < 1/k\}, k = 1, 2, 3, \ldots$ ดังนั้น $A_1 \cap A_2 \cap A_3 \cap \ldots$ เป็นเซทว่างเปล่า เนื่องจากว่าไม่มีจุดซึ่งเป็นของแต่ละเซท $A_1, A_2, A_3, \ldots$

ตัวอย่างที่ 10 ให้ A_1 และ A_2 เป็นเซทของจุดล้อมรอบด้วยสองวงกลมตัดกัน ดังนั้นเซท $A_1 \cup A_2$ กับ $A_1 \cap A_2$ **แทนที่ด้วยส่วนที่แรเงาในแผนภาพเวนน** ดังรูปที่ 1 และรูปที่ 2



$A \cup A$

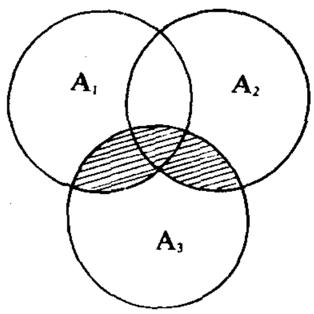
รูปที่ 1



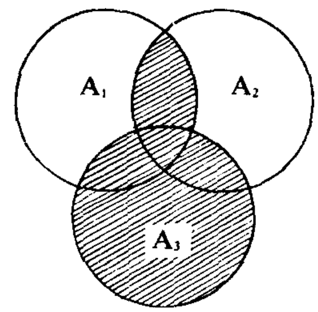
$A_1 \cap A_2$

รูปที่ 2

ตัวอย่างที่ 11 ให้ A_1, A_2 และ A_3 แทนเซทของจุดล้อมรอบด้วยสามวงกลมตัดกัน ดังนั้น เซท $(A_1 \cup A_2) \cap A_3$ และ $(A_1 \cap A_2) \cup A_3$ แสดงได้ดังรูปที่ 3 และ 4



รูปที่ 3 $(A_1 \cup A_2) \cap A_3$



รูปที่ 4 $(A_1 \cap A_2) \cup A_3$

นิยาม ให้ A และ B เป็นเซทสองเซท Cartesian product ของ A กับ B แสดงได้เป็น $A \times B$ หมายความว่า เซท $\{(a, b), a \in A, b \in B\}$ เป็นเซทของ ordered pairs ทั้งหมด ที่สมาชิกตัวแรกได้มาจาก A และสมาชิกตัวที่สองได้มาจาก B

ตัวอย่างที่ 1 2 ให้ $A = \{1, 2, 3\}$; $B = \{1, 2, 3, 4\}$ แล้ว

$A \times B = \{(1,1), (1,2), \ldots (1,4), (2,1), \ldots (2,4), (3,1), \ldots (3,4)\}$

สังเกตว่า โดยทั่ว ๆ ไป $A \times B \neq B \times A$

ถ้าหากว่า $A_1, A_2, \ldots A_n$ เป็นเซท n เซทแล้ว $A_1 \times A_2 \times \ldots \times A_n = \{(a_1, a_2, \ldots a_n), a_i \in A_i\}$

นั่นคือเซทของ ordered n-tuples ทั้งหมด

Sample Space sample space เป็นเซทของผลลัพธ์ที่อาจเป็นไปได้ทั้งหมด เราใช้สัญลักษณ์ S

ตัวอย่างที่ 1 ทอดลูกเต๋าหนึ่งลูกและสังเกตหมายเลขที่ปรากฏ $S = \{1, 2, 3, 4, 5, 6\}$

ตัวอย่างที่ 2 โยนเหรียญหนึ่งอันสี่ครั้ง และสังเกตจำนวนของหัวที่ได้

$S = \{0, 1, 2, 3, 4\}$

ตัวอย่างที่ 3 โรงงานผลิตสินค้าชนิดหนึ่ง และนับจำนวนสินค้าที่ผลิตเสียในระหว่าง 24 ชั่วโมง

$S = \{0, 1, 2, \ldots, N\}$ ในเมื่อ N เป็นจำนวนสูงสุดที่ผลิตใน 24 ชั่วโมง

ตัวอย่างที่ 4 นับจำนวนน๊อตเสียที่ประกอบปีกเครื่องบินจำนวนมาก

$S = \{0, 1, 2, \ldots, M\}$ ในเมื่อ M เป็นจำนวนน๊อตที่ประกอบปีกเครื่องบิน

ตัวอย่างที่ 5 โรงงานผลิตหลอดไฟ ต้องการทดสอบอายุของหลอดไฟ โดยเอาหลอดไฟไปเสียบปลั๊กจนกระทั่งหลอดไฟขาด และบันทึกจำนวนเวลาไว้

$S = \{t \mid t > 0\}$

ตัวอย่างที่ 6 กล่องใบหนึ่งมีหลอดไฟอยู่ 10 หลอด ใน 10 หลอดนี้มีหลอดเสียอยู่ 3 หลอด หยิบหลอดไฟครั้งละหลอด (หยิบแล้วไม่ใส่คืน) จนกระทั่งได้หลอดเสียหลอดสุดท้าย นับจำนวนหลอดทั้งหมดที่หยิบจากกล่อง

$S = \{3, 4, 5, 6, 7, 8, 9, 10\}$

ตัวอย่างที่ 7 ผลิตสินค้าจนกระทั่งผลิตสินค้าได้ไม่เสีย 10 ชิ้น นับจำนวนสินค้าที่ผลิตทั้งหมด

$S = \{10, 11, 12, \ldots\}$

ตัวอย่างที่ 8 วัดความแข็งแรงของท่อนเหล็กกล้า

$S = \{T \mid T \geq 0\}$

กฎของเซท ตามวัตถุประสงค์ทางคณิตศาสตร์ เซทมีพฤติกรรมตามสมมติฐานที่แน่นอน ซึ่งจะกล่าวต่อไป โดยให้เซท A, B และ C เป็นเซทย่อยทั้งหมดของ sample space S

กฎข้อที่ 1 Closure Laws

สำหรับแต่ละคู่ของเซท A กับ B จะมีเซท $A \cup B$ เซทเดียวเท่านั้น กับจะมีเซท $A \cap B$ เดียวเท่านั้น ของ sample space S

กฎข้อที่ 2 Commutative Laws

$$A \cup B = B \cup A \text{ และ } A \cap B = B \cap A$$

กฎข้อที่ 3 Associative Laws

$$(A \cup B) \cup C = A \cup (B \cup C)$$

และ

$$(A \cap B) \cap C = A \cap (B \cap C)$$

กฎข้อที่ 4 Distributive Laws

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

และ

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

กฎข้อที่ 5 Identity Laws

$A \cap S = A$ สำหรับแต่ละเซท A และจะมีเซท $\varnothing$ เท่านั้น **ที่ซึ่ง**

$A \cup \varnothing = A$ สำหรับแต่ละเซท A

กฎข้อที่ 6 Complementation Law

สำหรับแต่ละเซท A จะมีเซท $\bar{A}$ เท่านั้นที่ซึ่ง $A \cap \bar{A} = \varnothing$ และ $A \cup \bar{A} = S$

สังเกตว่า กฎข้อที่ห้าเราหมายถึงเซทว่างเปล่า และกฎข้อที่หกเราหมายถึง Complement $\bar{A}$ ของเซท A

กฎเหล่านี้ง่ายต่อการพิสูจน์ ถ้าหากว่าเราให้เซทเป็นที่รวมของจุดหรือของวัตถุอื่น ๆ

เหตุการณ์ เหตุการณ์หนึ่งหมายถึงเซทย่อยของ sample space S, sample space S ตัวของมันเองก็เป็นเหตุการณ์หนึ่ง ดังเช่น เซทว่างเปล่า $\varnothing$ ผลลัพธ์แต่ละผลลัพธ์ก็อาจเป็นเหตุการณ์หนึ่ง ดังตัวอย่าง

1. ทอดลูกเต๋าหนึ่งลูกหนึ่งครั้ง A_1 เป็นหมายเลขคู่ที่เกิดขึ้น นั่นคือ $A_1 = \{2, 4, 6\}$

2. โรงงานผลิตสินค้าชนิดหนึ่ง และนับจำนวนสินค้าที่ผลิตเสียในระหว่าง 24 ชั่วโมง ให้ A_2 เป็นจำนวนสินค้าที่ผลิตได้ไม่เสีย นั่นคือ $A_2 = \{0\}$

3. บันทึกอุณหภูมิติดต่อกันตลอด 24 ชั่วโมง ให้ x และ y เป็นอุณหภูมิต่ำสุด และสูงสุด ให้ A_3 เป็นอุณหภูมิสูงสุดที่มีค่ามากกว่าอุณหภูมิต่ำสุด 20° นั่นคือ $A_3 = \{(x, y) \mid y = x + 20\}$

1.4 ฟังก์ชันของเซท

ในแคลคูลัส ฟังก์ชัน คือ

$$f(x) = 2x \qquad -\infty < x < \infty$$

หรือ

$$g(x, y) = e^{-x-y} \qquad 0 < x < \infty, 0 < y < \infty$$
$$= 0 \text{ สำหรับค่าอื่น ๆ}$$

หรืออาจจะเป็น

$$h(x_1, x_2, \ldots, x_n) = 3x_1, x_2, \ldots, x_n \quad 0 \le x_i \le 1, i = 1, 2, \ldots, n$$
$$= 0 \text{ สำหรับค่าอื่น ๆ}$$

ค่าของ $f(x)$ ที่จุด $x = 1$ คือ $f(1) = 2$ ค่าของ $g(x, y)$ ที่จุด $(-1, 3)$ คือ $g(-1, 3) = e^{-2}$ ค่าของ $h(x_1, x_2, \ldots x_n)$ ที่จุด $(1, 1, \ldots 1)$ คือ 3 ฟังก์ชันเหล่านี้เรียกว่า ฟังก์ชันแบบจุด เพราะว่าคำนวณหาได้ที่จุดใน range หรือ space ของมิติที่ได้แสดงไว้ สำหรับฟังก์ชันของเซททั้งหมดของจุดโดยทั่ว ๆ ไป เรียกว่า ฟังก์ชันของเซท

ตัวอย่างที่ 1 ให้ A เป็นเซทหนึ่งใน space ที่เป็นหนึ่งมิติ และให้ Q(A) เป็นจำนวนของจุดใน A ซึ่งสมนัยกับเลขจำนวนเต็มบวก ดังนั้น Q(A) เป็นฟังก์ชันหนึ่งของเซท A ถ้า $A = \{x ; 0 < x < 5\}$ ดังนั้น $Q(A) = 4$ ถ้า $A = \{x ; x = -2, -1\}$ แล้ว $Q(A) = 0$; ถ้า $A = \{x ; -\infty < x < 6\}$ แล้ว $Q(A) = 5$

ตัวอย่างที่ 2 ให้ A เป็นเซทหนึ่งใน space สองมิติ และให้ Q(A) เป็นพื้นที่ของ A ถ้า A มีพื้นที่จำกัด อีกนัยหนึ่งให้ Q(A) ไม่จำกัด ถ้า $A = \{(x, y) ; x^2 + y^2 \le 1\}$ ดังนั้น $Q(A) = \pi$

ถ้า $A = \{(x, y) ; (x, y) = (0, 0), (1, 1), (0, 1)\}$ ดังนั้น $Q(A) = 0$ ถ้า $A = \{(x, y) ; 0 \le x, 0 \le y, x+y \le 1\}$ ดังนั้น $Q(A) = 1/2$

ตัวอย่างที่ 3 ให้ A เป็นเซทหนึ่งใน space สามมิติ และให้ Q(A) เป็นปริมาตรของ A ถ้า A มีปริมาตรที่จำกัด อีกนัยหนึ่งให้ Q(A) ไม่จำกัด ถ้า $A = \{(x, y, z) ; 0 \le x \le 2, 0 \le y \le 1, 0 \le z \le 3\}$ ดังนั้น $Q(A) = 6$ ถ้า $A = \{(x, y, z) ; x^2 + y^2 + z^2 \ge 1\}$ ดังนั้น Q(A) ไม่จำกัด

ต่อไปนี้จะเสนอแนะสัญลักษณ์

$$\int_A f(x)\,dx$$

หมายถึง integral ธรรมดาของ f (x) ตลอดทั้งหมดของเซท A ในหนึ่งมิติ สัญลักษณ์

$$\int_A \int g(x, y)\,dx\,dy$$

จะหมายถึง Riemann integral ของ g(x, y) ตลอดทั้งหมดของเซท A ในสองมิติและต่อ ๆ ไป การเลือกฟังก์ชัน f (x) และ g (x, y) ต้องระมัดระวัง เพราะอาจหาค่า integral ไม่ได้ ในทำนองเดียวกัน สัญลักษณ์

$$\sum_A f(x)$$

จะหมายถึง ผลรวมตลอดทั้งหมดของ $x \in A$ สัญลักษณ์

$$\sum_A \sum g(x, y)$$

จะหมายถึง ผลรวมตลอดทั้งหมดของ $(x, y) \in A$ และต่อ ๆ ไป

ตัวอย่างที่ 4 ให้ A เป็นเซทหนึ่งใน space หนึ่งมิติ และให้

$$Q(A) = \sum_A f(x) \text{ ในเมื่อ}$$

$$f(x) = (1/2)^x, \; x = 1, 2, 3, \ldots$$
$$= 0 \text{ สำหรับค่าอื่น ๆ}$$

ถ้า $A = \{x ; 1 \le x \le 3\}$ แล้ว

$$Q(A) = (1/2) + (1/2)^2 + (1/2)^3 = 7/8$$

ตัวอย่างที่ 5 ให้ $Q(A) = \sum_{A} f(x)$ ในเมื่อ

$$f(x) = p^x (1-p)^{1-x} \qquad x = 0, 1$$
$$= 0 \quad \text{สำหรับค่าอื่น ๆ}$$

ถ้า $A = \{x ; x = 0\}$ แล้ว

$$Q(A) = \sum_{x=0}^{0} p^x (1-p)^{1-x} = 1 - p$$

ถ้า $A = \{x ; 1 \leq x \leq 2\}$ แล้ว $Q(A) = f(1) = p$

ตัวอย่างที่ 6 ให้ A เป็นเซทหนึ่งมิติ และให้

$$Q(A) = \int_A e^{-x}\,dx$$

ถ้า $A = \{x ; 0 \leq x \leq \infty\}$ แล้ว

$$Q(A) = \int_0^{\infty} e^{-x}\,dx = 1$$

ถ้า $A = \{x ; 1 \leq x \leq 2\}$ แล้ว

$$Q(A) = \int_1^2 e^{-x}\,dx \doteq e^{-1} - e^{-2}$$

ถ้า $A_1 = \{x ; 0 \leq x \leq 1\}$ และ $A_2 = \{x ; 1 < x \leq 3\}$ แล้ว $A_1 \cup A_2 = \{x ; 0 \leq x \leq 3\}$

$$Q(A_1 \cup A_2) = \int_0^3 e^{-x}\,dx$$
$$= \int_0^1 e^{-x}\,dx + \int^3 e^{-x}\,dx$$
$$= Q(A_1) + Q(A_2)$$

ถ้า $A = A_1 \cup A_2$ ในเมื่อ $A_1 = \{x ; 0 \leq x \leq 2\}$ และ $A_2 = \{x ; 1 \leq x \leq 3\}$ แล้ว

$$Q(A) = Q(A_1 \cup A_2) = \int_0^3 e^{-x}\,dx$$

$$= \int_0^2 e^{-x}\,dx + \int_1^3 e^{-x}\,dx - \int_1^2 e^{-x}\,dx$$

$$= Q(A_1) + Q(A_2) - Q(A_1 \cap A_2)$$

ตัวอย่างที่ 7 ให้ A เป็นเซทหนึ่งใน space n- มิติ และให้

$$Q(A) = \int \cdots \int_A dx_1\, dx_2 \ldots dx_n$$

ถ้า $A = \{(x_1, x_2, \ldots, x_n)\,;\, 0 \le x_1 \le x_2 \le \ldots \le x_n \le 1\}$ แล้ว

$$Q(A) = \int_0^1 \int_0^{x_n} \cdots \int_0^{x_3} \int_0^{x_2} dx_1\, dx_2 \ldots dx_{n-1}\, dx_n$$

$$= \int_0^1 \int_0^{x_n} \cdots \int_0^{x_3} x_2\, dx_2\, dx_3 \ldots dx_{n-1}\, dx_n$$

$$= \int_0^1 \int_0^{x_n} \cdots \int_0^{x_4} \frac{x_2^2}{2!} \Big|_0^{x_3} dx_3 \ldots dx_{n-1}\, dx_n$$

$$= \int_0^1 \int_0^{x_n} \cdots \int_0^{x_5} \int_0^{x_4} \frac{x_3^2}{2!}\, dx_3\, dx_4 \ldots dx_{n-1}\, dx_n$$

$$= \ldots\ldots\ldots\ldots\ldots\ldots$$

$$= \ldots\ldots\ldots\ldots\ldots\ldots$$

$$= \int_0^1 \int_0^{x_n} \frac{x_{n-1}^{n-1}}{(n-2)!} dx_{n-1}\, dx_n$$

$$= \int_0^1 \frac{x_n^{n-1}}{(n-1)!}\, dx_n$$

$$= \frac{x_n^n}{n!} \Big|_0^1$$

$$= \frac{1}{n!}$$ ในเมื่อ $n! = n(n-1)(n-2) \ldots 3 \cdot 2 \cdot 1$

1.5 ทฤษฎีความน่าจะเป็น

การทดลองเชิงสุ่มเป็นจุดเริ่มต้นของการศึกษาทฤษฎีความน่าจะเป็น ดังนั้น ทฤษฎี-ความน่าจะเป็นจึงหมายถึงศาสตร์อย่างหนึ่ง และศาสตร์นั้นว่าด้วยความไม่แน่นอน ซึ่งเป็นผลลัพธ์มาจากการทดลองเชิงสุ่ม โดยหวังว่าผลลัพธ์จากทฤษฎีความน่าจะเป็นนี้ จะช่วยให้คนเราเอาชนะความไม่แน่นอนจากการทดลองเชิงสุ่ม

นิยาม ให้ E เป็นการทดลองหนึ่ง S เป็น sample space ที่สมนัยกับ E A เป็นเหตุการณ์หนึ่งของเลขจำนวนจริง P(A) เรียกว่า ความน่าจะเป็นของเหตุการณ์ซึ่งมีคุณสมบัติดังนี้

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. ถ้า A และ B เป็นเหตุการณ์ที่ mutually exclusive กันแล้ว
   $P(A \cup B) = P(A) + P(B)$
4. ถ้าหากว่า $A_1, A_2, \ldots, A_n, \ldots$ เป็นเหตุการณ์ mutually exclusive ของแต่ละคู่ซึ่งกันและกันแล้ว
   $P(\cup_{i=1}^{\infty} A_i) = P(A_1) + P(A_2) + \ldots + P(A_n) + \ldots$

จากคุณสมบัติข้อสาม หากเหตุการณ์มีจำนวน n เหตุการณ์

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

ความน่าจะเป็นของผลลัพธ์ในความหมายของคณิตศาสตร์ จะได้ว่า

$P(O_i)$ เป็นตัวเลขของผลลัพธ์ O_i ที่อยู่ใน sample space S

$P(O_i)$ เป็นฟังก์ชันของจุดที่เป็นเลขจำนวนจริง ซึ่งนิยามใน domain ซึ่งเป็น sample space และมีพิสัย (range) เป็นเลขจำนวนจริง

ทฤษฎีที่ 1 ถ้า $\varnothing$ เป็นเซทว่างเปล่าแล้ว $P(\varnothing) = 0$

พิสูจน์ เราอาจเขียนสำหรับเหตุการณ์หนึ่งเหตุการณ์ใด A,

$A = A \cup \varnothing$ เนื่องจาก A กับ $\varnothing$ เป็น mutually exclusive

จากคุณสมบัติข้อ 3 ค่า $P(A) = P(A \cup \varnothing) = P(A) + P(\varnothing)$

จากนี้ เราก็สามารถสรุปทฤษฎีได้

ข้อสังเกต ผลกลับของทฤษฎีข้างต้นไม่เป็นจริง นั่นคือ
ถ้า $P(A) = 0$ โดยทั่ว ๆ ไปเราไม่สามารถสรุปว่า $A = \varnothing$ สำหรับมีหลาย ๆ สภาวะ ในที่ซึ่งเรากำหนดความน่าจะเป็นเป็นศูนย์ ต่อเหตุการณ์ที่สามารถเกิดขึ้น

ทฤษฎีที่ 2 ถ้า $\overline{A}$ เป็นเหตุการณ์ที่ Complementary ของเหตุการณ์ A แล้ว

$$P(A) = 1 - P(\overline{A})$$

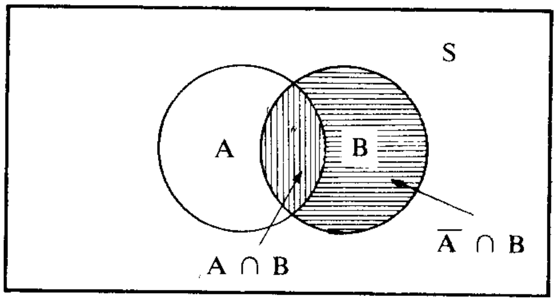
พิสูจน์ เราอาจเขียน $S = A \cup \overline{A}$ และใช้คุณสมบัติข้อ 2 และ 3

เราได้

$$\begin{aligned} P(S) &= P(A \cup \overline{A}) \\ 1 &= P(A) + P(\overline{A}) \\ P(A) &= 1 - P(\overline{A}) \end{aligned}$$

ทฤษฎีที่ 3 ถ้าหากว่า A กับ B เป็นสองเหตุการณ์ใด ๆ แล้ว

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



รูปที่ 5

พิสูจน์ การพิสูจน์นี้ ต้องแยกด้วยเหตุการณ์ $A \cup B$ กับ B เป็นเหตุการณ์ที่ mutually exclusive กัน เราใช้คุณสมบัติข้อ 3 ดูรูปที่ 5

ดังนั้น เราเขียนได้

$$\begin{aligned} A \cup B &= A \cup (B \cap \overline{A}) \\ B &= (A \cap B) \cup (B \cap \overline{A}) \end{aligned}$$

จากนี้

$$\begin{aligned} P(A \cup B) &= P(A) + P(B \cap \overline{A}) \\ P(B) &= P(A \cap B) + P(B \cap \overline{A}) \end{aligned}$$

เอาสองสมการนี้ลบกันได้

$$P(A \cup B) - P(B) = P(A) - P(A \cap B)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

ทฤษฎีนี้ขยายคุณสมบัติข้อที่ 3 หาก $A \cap B = \varnothing$ เราจะได้ข้อความของคุณสมบัติข้อที่ 3

ทฤษฎีที่ 4 ถ้าหากว่า A, B และ C เป็นสามเหตุการณ์แล้ว

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$

พิสูจน์ การพิสูจน์นี้ให้เขียน $A \cup B \cup C$ เป็น $(A \cup B) \cup C$ และใช้ผลลัพธ์ของทฤษฎีข้างต้น ส่วนรายละเอียดให้นักศึกษาไปพิสูจน์เอง

ข้อสังเกต สำหรับกรณีที่มีเหตุการณ์มาก ๆ อย่างเช่น ให้ $A_1, A_2, \ldots, A_k$ เป็น k เหตุการณ์แล้ว

$$P(A_1 \cup A_2 \cup \ldots \cup A_k) = \sum_{i=1}^{k} P(A_i) - \sum_{i<j=2}^{k} P(A_i \cap A_j) + \sum_{i<j<r=3}^{k} P(A_i \cap A_j \cap A_r) + \ldots + (-1)^{k-1} P(A_1 \cap A_2 \cap \ldots \cap A_k)$$

ผลลัพธ์นี้ใช้หลักการพิสูจน์ทางคณิตศาสตร์

ทฤษฎี ถ้าหากว่า $A \subset B$ แล้ว $P(A) \leq P(B)$

พิสูจน์ เราอาจแยก B เป็นสองเหตุการณ์ที่เป็น mutually exclusive ได้ดังนี้

$B = A \cup (B \cap \overline{A})$ จากนี้ $P(B) = P(A) + P(B \cap \overline{A}) \geq P(A)$

เนื่องจากว่า $P(B \cap \overline{A}) \geq 0$ จากคุณสมบัติข้อที่ 1

ทฤษฎี ถ้าหากว่า $A \subset S$ แล้ว

$$0 \leq P(A) \leq 1$$

พิสูจน์ เนื่องจากว่า $\varnothing \subset A \subset S$ จากทฤษฎีข้างต้น เราได้

$$P(\varnothing) \leq P(A) \leq P(S)$$
$$0 \leq P(A) \leq 1$$

1.6 ตัวแปรเชิงสุ่ม

สมมติว่า ผลลัพธ์ (outcome) ของการทดลองเชิงสุ่ม สามารถแสดงออกเป็นจำนวนเดี่ยว ๆ แล้ว sample space S สามารถแทนได้ด้วยเซทของจุดบนเส้นโดยตรง ถ้าหากว่า เราแสดงผลลัพธ์ด้วยสัญลักษณ์ X เราเรียก X ว่า ตัวแปรเชิงสุ่ม ถ้าหากว่า A เป็นเซทย่อยของ S เราควรเขียน

$$P(A) = \text{ความน่าจะเป็นที่ } X \in A = P(X \in A)$$

ถ้าหากว่า ผลลัพธ์ของการทดลองเชิงสุ่มแสดงออกเป็น ordered pair ของตัวเลข เราก็สามารถแทนผลลัพธ์นี้ได้ด้วยสองตัวแปรเชิงสุ่ม X กับ Y sample space คือ เซทของจุดใน space สองมิติ ถ้าหากว่า A เป็นเซทย่อยของ S เราเขียนได้เป็น

$$P(A) = \text{ความน่าจะเป็นที่ } (X, Y) \in A = P[(X, Y) \in A]$$

โดยทั่ว ๆ ไป ถ้าหากว่าผลลัพธ์ของการทดลองเชิงสุ่มแสดงออกเป็น n ordered number ผลลัพธ์นั้นก็แทนได้ด้วย n ตัวแปรเชิงสุ่ม $X_1, X_2, X_3, \ldots, X_n$ sample space S คือเซทของจุดใน space n มิติ และ A เป็นเซทย่อยของ S เราเขียนได้เป็น

$$P(A) = P[(X_1, X_2, \ldots, X_n) \in A]$$

ถ้าหากว่า เซท $A_1, A_2, A_3, \ldots$ ไม่มีจุดร่วมกันแล้ว เซท $A_1, A_2, \ldots$ เรียกว่า mutually disjoint sets หรือ mutually exclusive events ดังตัวอย่าง ถ้าหากว่า S เป็นหนึ่งมิติ เราได้

$$P(X \in A_1 \cup A_2 \cup \ldots) = P(X \in A_1) + P(X \in A_2) + \ldots$$

Probability set function P(A) ของหนึ่งตัวแปรเชิงสุ่ม หรือมากกว่า เรียกว่า ฟังก์ชัน

การแจกแจงความน่าจะเป็น เพราะว่ามันบอกความน่าจะเป็นทั้งหมดว่า มีการแจกแจงคลุมเซท ต่าง ๆ ใน sample space ว่าเป็นอย่างไร

ตัวอย่างที่ 1 ให้ probability set function P(A) กำหนดได้โดย

$$P(A) = \int_A f(x)\,dx \text{ ในเมื่อ } f(x) = \frac{3x^2}{8}$$

$x \in S = \{x ; 0 < x < 2\}$ เนื่องจากว่า S เป็นเซทใน space หนึ่งมิติ ที่มีตัวแปรเชิงสุ่ม X ให้ $A_1 = \{x ; 0 < x < 1/2\}$ กับ $A_2 = \{x ; 1 < x < 2\}$ เป็นเซทของ S แล้ว

$$P(A_1) = P(X \in A_1) = \int_{A_1} f(x)\,dx = \int_0^{1/2} \frac{3x^2}{8}\,dx = \frac{x^3}{8}\Big|_0^{1/2} = \frac{1}{64}$$

และ $$P(A_2) = P(X \in A_2) = \int_{A_2}^{1} f(x)\,dx = \int_1^2 \frac{3x^2}{8}\,dx = \frac{x^3}{8}\Big|_1^2 = \frac{7}{8}$$

ถ้าหากว่า เราต้องการคำนวณ $P(A_1 \cup A_2)$ ในกรณี $A_1 \cap A_2 = 0$ เราจะได้

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) = \frac{57}{64}$$

ตัวอย่างที่ 2 ให้ $S = \{(x, y) ; 0 < x < y < 1\}$ นี้เป็นเซทใน space สองมิติ เกี่ยวข้องกับสองตัวแปรเชิงสุ่ม X กับ Y ให้ probability set function นิยามได้

$$P(A) = \int_A\int 2dx\,dy$$

ถ้า $A_1 = \{(x, y) ; 1/2 < x < y < 1\}$ แล้ว

$$P(A_1) = P[(X, Y) \in A_1] = \int_{1/2}^{1}\int_{1/2}^{y} 2dx\,dy$$

$$= \int_{1/2}^{1} 2x\Big|_{1/2}^{y} dy = \int_{1/2}^{1} (2y - 1)\,dy$$

$$= 2\int_{1/2}^{1} (y - 1/2)\, d(y - 1/2) = (y - 1/2)^2 \Big|_{1/2}^{y} = 1/4$$

ถ้า $A_2 = \{(x, y) ; x < y < 1, 0 < x \leq 1/2\}$ แล้ว $A_2 = \overline{A_1}$

$$P(A_2) = P[(X, Y) \in A_2] = P(\overline{A_1}) = 1 - P(A_1) = 3/4$$

ตัวอย่างที่ 3 ให้ $S = \{(x, y) ; (x, y) = (1, 1), \ldots, (1, 6), (2, 1), \ldots, (2, 6), \ldots, (6, 6)\}$ นั่นคือ S มี 36 สมาชิก ถ้า $A \subset S$ ให้ probability set function กำหนดได้เป็น

$$P(A) = \sum_{A}\sum f(x, y) \text{ ในเมื่อ } f(x, y) = \frac{1}{36}, (x, y) \in S$$

ถ้า $A_1 = \{(x, y) ; (x, y) = (1, 6), (2, 5), (3, 4)\}$ และ $A_2 = \{(x, y) ; (x, y) = (3, 4), (5, 6)\}$ แล้ว

$$P(A_1) = f(1, 6) + f(2, 5) + f(3, 4) = 3/36$$

และ $$P(A_2) = f(3, 4) + f(5, 6) = \frac{2}{36}$$

$$P(A_1 \cap A_2) = f(3, 4) = \frac{1}{36}$$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = \frac{4}{36}$$

1.7 ฟังก์ชันความหนาแน่นน่าจะเป็น

นิยาม ฟังก์ชันของตัวแปรเชิงสุ่ม X เป็นฟังก์ชันความหนาแน่น (pdf) จะต้องสอดคล้องเงื่อนไขดังต่อไปนี้

ในกรณีตัวแปรเชิงสุ่ม X เป็นชนิดไม่ต่อเนื่อง

ก. $f(x) > 0$ สำหรับ X ทั้งหมด

ข. $\sum_{S} f(x) = 1$

ค. $P(A) = P(X \in A) = \sum_{A} f(x)$

ตัวอย่างที่ 1 ให้ X เป็นตัวแปรเชิงสุ่มชนิดไม่ต่อเนื่อง ซึ่งมี sample space

$S = \{x ; x = 0, 1, 2, 3, 4\}$ ให้

$$P(A) = \sum_{A} f(x)$$

ในเมื่อ

$$f(x) = \frac{4!}{x!\,(4-x)!}(1/2)^4 \ , \quad x \in S$$

และ $0! = 1$ แล้ว ถ้า $A = \{x ; x = 0, 1\}$ เราได้

$$P(X \in A) = \frac{4!}{0!\,4!}(1/2)^4 + \frac{4!}{1!\,3!}(1/2)^4$$

$$= \frac{5}{16}$$

ตัวอย่างที่ 2 ให้ $S = \{x ; x = 1, 2, 3, \ldots\}$ และให้

$$f(x) = (1/2)^x ; x \in S$$

ถ้า X เป็นตัวแปรเชิงสุ่มชนิดไม่ต่อเนื่อง เพื่อว่า

$$P(X \in A) = \sum_{A} f(x)$$

แล้วถ้า $A = \{x ; x = 1, 3, 5, 7, \ldots\}$ เราได้

$$P(X \in A) = (1/2) + (1/2)^3 + (1/2)^5 + \ldots$$

$$= \frac{2}{3}$$

ในกรณีตัวแปรเชิงสุ่ม X เป็นชนิดต่อเนื่อง

ก. $f(x) > 0$ สำหรับ x ทั้งหมด

ข. $\int_S f(x)\,dx = 1$

ค. $P(A) = P(X \in A) = \int_A f(x)\,dx$

ตัวอย่างที่ 3 ตัวแปรเชิงสุ่ม X มีฟังก์ชันน่าจะเป็นดังรูปต่อไปนี้ ในเมื่อ k เป็นค่าคงที่

$$f(x) = \begin{cases} k & x = 0 \\ 2k & x = 1 \\ 3k & x = 2 \\ 0 & \text{สำหรับค่าอื่น ๆ} \end{cases}$$

ก. คำนวณค่า k

ข. คำนวณหา $P(X < 2)$, $P(X \leq 2)$, $P(0 < X < 2)$

ค. ค่าของ k ที่มีค่าน้อยที่สุดสำหรับซึ่ง $P(X \leq k) > 1/2$

วิธีทำ ก. $f(x_i) > 0$ สำหรับ $i = 0, 1, 2$ และ $k > 0$

เนื่องจากว่า $\sum_{i=0}^{2} f(x_i) = 1$

$$k + 2k + 3k = 1$$

$$6k = 1$$

$$k = 1/6$$

ข.

$$P(X < 2) = P(X = 0) + P(X = 1)$$

$$= k + 2k = 1/6 + 2/6 = 1/2$$

$$P(X \leq 2) = P(X < 2) + P(X = 2)$$

$$= \frac{1}{2} + 3k = \frac{1}{2} + \frac{3}{6} = 1$$

$$P(0 < X < 2) = P(X = 1) = 2k = \frac{2}{6} = \frac{1}{3}$$

ค. $P(X \leq k) > 1/2$ โดยการทดลอง ถ้าหาก $k = 1$

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$= \frac{1}{6} + \frac{1}{3} = 1/2$$

มีค่าน้อยกว่าโจทย์กำหนดให้ คือ มากกว่า 1/2

ถ้าหากว่า $k = 2$

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$= 1$$

ซึ่งมากกว่า 1/2 ดังนั้น $k = 2$

ตัวอย่างที่ 4 ให้ $S = \{x ; 0 < x < \infty\}$ และให้

$f(x) = e^{-x}, x \in S$

ถ้า X เป็นตัวแปรเชิงสุ่มชนิดต่อเนื่อง

$P(X \in A) = \int_A e^{-x}\, dx$

เรามี $A = \{x ; 0 < x < 1\}$

$P(X \in A) = \int_0^1 e^{-x}\, dx = 1 - e^{-1}$

ตัวอย่างที่ 5 ให้ X เป็นตัวแปรเชิงสุ่มชนิดต่อเนื่องมี sample space $S = \{x ; 0 < x < 1\}$

ให้ probability set function กำหนดได้เป็น

$$P(A) = \int_A f(x)\, dx$$

ในเมื่อ

$$f(x) = Cx^2 \qquad x \in S$$

เนื่องจากว่า P(A) เป็น probability set function $P(S) = 1$
ดังนั้น ค่าคงที่ C คำนวณหาได้โดย

$$\int_0^1 Cx^2\, dx = 1$$

$$\frac{Cx^3}{3}\Big|_0^1 = 1 \quad ; \qquad C = 3$$

สัญลักษณ์ของ pdf ของหนึ่งตัวแปรเชิงสุ่ม X สามารถขยายออกไปถึง pdf ของสองตัวแปรเชิงสุ่มหรือมากกว่าภายใต้ข้อจำกัดของ sample space S แทนฟังก์ชัน $f > 0$ บน S ในกรณีสองตัวแปรเชิงสุ่ม X กับ Y เป็นชนิดไม่ต่อเนื่อง probability set function P(A), $A \subset S$ กำหนดได้โดย

$$P(A) = P[(X, Y) \in A] = \sum_A \sum f(x, y)$$

ในกรณีเป็นชนิดต่อเนื่อง

$$P(A) = P[(X, Y) \in A] = \int_A \int f(x, y)\, dx\, dy$$

ในแต่ละกรณี f เรียกว่า pdf ของสองตัวแปรเชิงสุ่ม X กับ Y และ P(S) = 1 โดยทั่ว ๆ ไป n ตัวแปรเชิงสุ่ม $X_1, X_2, \ldots, X_n$ เป็นชนิดไม่ต่อเนื่องหรือต่อเนื่อง และมี probability set function P(A), $A \subset S$ กำหนดได้โดย

$$P(A) = P[(X_1, X_2, \ldots, X_n) \in A] = \Sigma \underset{A}{\Sigma} \ldots \Sigma f(x_1, x_2, \ldots, x_n)$$

หรือ $$P(A) = P[(X_1, X_2, \ldots, X_n) \in A] = \int\int\underset{A}{\int}\ldots\int f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n$$

แนวทางการศึกษาคือ ฟังก์ชัน f ไม่ว่าหนึ่งตัวแปรหรือมากกว่า จะสอดคล้องเงื่อนไขของการเป็น pdf ถ้าหากว่า $f > 0$ บน sample space S ถ้าหากว่า อินทิกรัลหรือผลรวมของฟังก์ชันตลอดทั้งหมดของ S มีค่าเป็นหนึ่ง

เพื่อพิจารณาให้ง่ายขึ้น เราสามารถเขียนได้จากตัวอย่าง เช่น sample space ชนิดต่อเนื่องของตัวแปรเชิงสุ่ม X เป็น $S = \{x\,;\, 0 < x < \infty\}$ และ pdf ของ X เป็น e^{-x} $x \in S$ เป็น

$$f(x) = e^{-x} \quad , \quad 0 < x < \infty$$
$$= 0 \quad \text{สำหรับค่าอื่น ๆ}$$

และ $\int_S f(x)\, dx$ แทนได้โดย $\int_{-\infty}^{\infty} f(x)\, dx$

ในทำนองเดียวกัน เราอาจขยายคำจำกัดความของ pdf f(x, y) ทั้งหมด ของระนาบ xy หรือ pdf f(x, y, z) ทั้งหมดของ space สามมิติและต่อ ๆ ไป

$$\int_S\int f(x, y)\, dx\, dy \quad \text{แทนด้วย} \quad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy$$

ในทำนองเดียวกัน คำจำกัดความของ pdf ของชนิดไม่ต่อเนื่อง

กรณีหนึ่งตัวแปรเชิงสุ่ม

$$\sum_S f(x) \text{ แทนด้วย } \sum_x f(x)$$

กรณีสองตัวแปรเชิงสุ่ม

$$\Sigma_S\Sigma f(x, y) \text{ แทนด้วย } \sum_y \sum_x f(x, y)$$

และกรณีต่อ ๆ ไป

ถ้า f(x) เป็น pdf ชนิดต่อเนื่องของตัวแปรเชิงสุ่ม X และถ้า A เป็นเซท $\{x ; a < x < b\}$ แล้ว $P(A) = P(X \in A)$ สามารถเขียนได้เป็น

$$P(a < X < b) = \int_a^b f(x)\,dx$$

ถ้า $A = \{x ; x = a\}$ แล้ว

$$P(A) = P(X \in A) = P(X = a) = \int_a^a f(x)\,dx = 0$$

ดังนั้น ถ้า X เป็นตัวแปรเชิงสุ่มของชนิดต่อเนื่อง ความน่าจะเป็นของทุก ๆ เซทประกอบด้วยจุดเดียว มีค่าเป็นศูนย์ จากความจริงนี้ สามารถทำให้เราเขียน

$$P(a < X < b) = P(a \le X \le b)$$

ดังเช่น

$$f(x) = e^{-x} \qquad 0 < x < \infty$$
$$= 0 \qquad \text{สำหรับค่าอื่น ๆ}$$

สามารถเขียนได้เป็น

$$f(x) = e^{-x} \qquad 0 \le x < \infty$$

$$= 0 \quad \text{สำหรับค่าอื่น ๆ}$$

โดยปราศจากการเปลี่ยนค่าใด ๆ สังเกตว่าสองฟังก์ชันเหล่านี้ ต่างกันที่ $X=0$ กับ $P(X=0)=0$ เท่านั้น ถ้าหากว่า สอง pdf ของตัวแปรเชิงสุ่มชนิดต่อเนื่องต่างกันเกี่ยวกับเซทหนึ่งมีความน่าจะเป็นเป็นศูนย์เท่านั้น สอง probability set functions ที่สมนัยกัน จะเหมือนกันอย่างแน่นอน ส่วน pdf ชนิดไม่ต่อเนื่องของตัวแปรเชิงสุ่มอาจไม่เปลี่ยนที่จุดใดจุดหนึ่ง เนื่องจากว่าการเปลี่ยนใน pdf เช่นนั้นจะเปลี่ยนการแจกแจงของความน่าจะเป็น

1.8 ฟังก์ชันการแจกแจง

ในตัวแปรเชิงสุ่ม X มี probability set function P(A) ในเมื่อ A เป็นเซทหนึ่งมิติ ให้ x เป็นเลขจำนวนจริง และเซท A เป็นเซทจาก $-\infty$ ถึง x รวมทั้งจุด x ด้วย สำหรับเซท A ทั้งหมด เราได้ $P(A) = P(X \in A) = P(X \le x)$ ความน่าจะเป็นนี้ขึ้นอยู่กับจุด x นั่นคือ ความน่าจะเป็นนี้ คือ ฟังก์ชันของจุด x ฟังก์ชันนี้แสดงได้โดยสัญลักษณ์ $F(x) = P(X \le x)$

ฟังก์ชัน $F(x)$ เรียกว่า ฟังก์ชันการแจกแจงหรือฟังก์ชันการแจกแจงสะสมของตัวแปรเชิงสุ่ม X

นิยาม ให้ X เป็นตัวแปรเชิงสุ่มชนิดไม่ต่อเนื่อง หรือต่อเนื่อง เรากำหนด F เป็นฟังก์ชันการแจกแจงของตัวแปรเชิงสุ่ม X (cdf) ในเมื่อ $F(x) = P(X \leq x)$

ทฤษฎีที่ 1 ก. ถ้าหากว่า X เป็นตัวแปรเชิงสุ่มชนิดไม่ต่อเนื่อง

$$F(x) = \sum_j f(x_j)$$

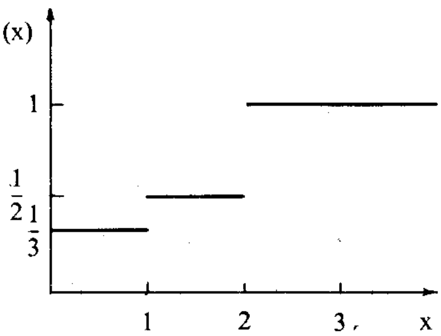
ในเมื่อผลรวมได้มาจากดัชนี j ทั้งหมดที่สอดคล้อง $x_j \leq x$

ข. ถ้าหากว่า X เป็นตัวแปรเชิงสุ่มชนิดต่อเนื่อง

$$F(x) = \int_{-\infty}^{x} f(w)\,dw$$

ตัวอย่างที่ 1 สมมติว่าตัวแปรเชิงสุ่ม X มีสามค่า 0, 1 และ 2 พร้อมด้วยความน่าจะเป็น 1/3, 1/6 และ 1/2 ตามลำดับแล้ว

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/3 & 0 \leq x < 1 \\ 1/2 & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

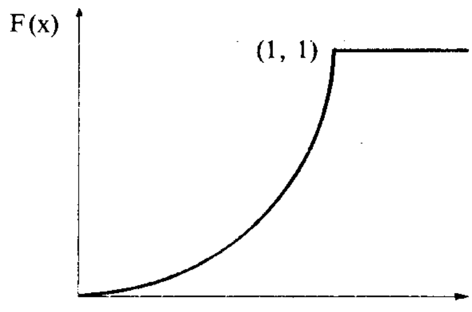


ตัวอย่างที่ 2 สมมติว่า X เป็นตัวแปรเชิงสุ่มชนิดต่อเนื่อง พร้อมด้วย pdf

$$f(x) = \begin{cases} 2x & 0 < x < 1 \\ 0 & \text{สำหรับค่าอื่น ๆ} \end{cases}$$

ดังนั้น cdf F กำหนดได้โดย

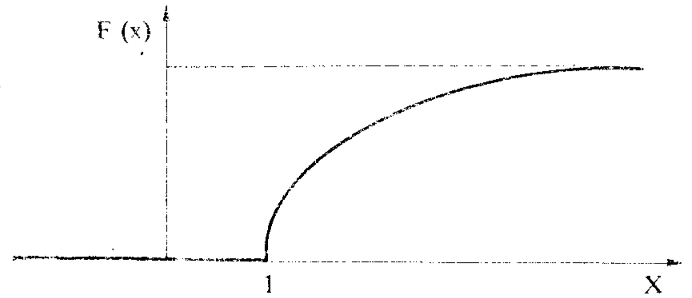
$$F(x) = \begin{cases} 0 & x \leq 0 \\ \int_0^x 2s\,ds = x^2 & 0 < x \leq 1 \\ 1 & x > 1 \end{cases}$$



ตัวอย่างที่ ๒ ให้ตัวแปรเชิงสุ่ม X ชนิดต่อเนื่องมี pdf $f(x) = \frac{2}{x^3}$, $1 < x < \infty$ ศูนย์สำหรับค่าอื่น ๆ ฟังก์ชันการแจกแจงของ X คือ

$$
\begin{aligned}
F(x) &= \int_{-\infty}^{x} 0\,dx = 0 &, x < 1 \\
&= \int_{1}^{x} \frac{2}{w^3}\,dw = 1 - \frac{1}{x^2} &, 1 \le x
\end{aligned}
$$

กราฟของฟังก์ชันการแจกแจงนี้เขียนได้



F(x) เป็นฟังก์ชันที่ต่อเนื่องสำหรับเลขจำนวนจริง x ทั้งหมด โดยเฉพาะ F(x) ทุก ๆ ที่ต่อเนื่องไปทางขวา นอกจากนั้น อนุพันธ์ (derivative) ของ F(x) เทียบกับ x หาค่าได้ที่จุดทั้งหมดนอกจากที่ $x = 1$ ดังนั้น pdf ของ X ถูกกำหนดได้ด้วยอนุพันธ์นี้นอกจากที่ $x = 1$ เนื่องจากว่า เซท $A = \{x ; x = 1\}$ เป็นเซทของ $P(A) = 0$ เรามีความอิสระที่จะกำหนด pdf ที่จุด $x = 1$ ในลักษณะใดลักษณะหนึ่งที่พอใจ วิธีหนึ่งที่จะทำนี้คือ ต้องเขียน $f(x) = \frac{2}{x^3}$, $1 < x < \infty$ ศูนย์สำหรับค่าอื่น ๆ

คุณสมบัติของฟังก์ชันการแจกแจง F(x) เราจะไม่บังคับ X ว่าจะเป็นตัวแปรเชิงสุ่มชนิดต่อเนื่องหรือไม่ต่อเนื่อง ดังนี้

ก. $0 \le F(x) \le 1$ เพราะว่า $0 \le P(X \le x) \le 1$

ข. F(x) เป็นฟังก์ชันเพิ่มขึ้นของค่า x สำหรับถ้า $x' \le x''$ แล้ว

$$\{x ; x \le x''\} = \{x ; x \le x'\} \cup \{x ; x' < x \le x''\}$$

และ

$$P(X \le x'') = P(X \le x') + P(x' < X \le x'')$$

นั่นคือ

$$F(x'') - F(x') = P(x' < X \le x'') \ge 0$$

ค. $F(\infty) = 1$ และ $F(-\infty) = 0$ เพราะว่า เซท $\{x ; x \le \infty\}$ คือ space หนึ่งมิติทั้งหมด และเซท $\{x ; x \le -\infty\}$ เป็นเซทว่างเปล่า หรือ

$$F(-\infty) = \lim_{x \to -\infty} \int_{-\infty}^{x} f(w)\,dw = 0$$

$$F(\infty) = \lim_{x \to \infty} \int_{-\infty}^{x} f(w)\,dw = 1$$

จากคุณสมบัติข้อ (ข) ถ้า $a < b$ แล้ว

$$P(a < X \le b) = F(b) - F(a)$$

และสมมติว่าเราต้องการใช้ $F(x)$ เพื่อคำนวณความน่าจะเป็น $P(X = b)$ ให้เราพิจารณาพร้อมด้วย $h > 0$

$$\lim_{h \to 0} P(b - h < X \le b) = \lim_{h \to 0} [F(b) - F(b-h)]$$

$\lim_{h \to 0} P(b - h < X \le b)$ ควรหาค่าได้ และเท่ากับ $P(X = b)$ เพราะว่า ขณะที่ h เข้าใกล้ศูนย์ limit ของเซท $\{x ; b - h < X \le b\}$ เป็นเซทที่มีจุดเดียว $X = b$ และเป็นความจริงที่ limit นี้คือ $P(X = b)$ เป็นทฤษฎีที่เรายกเว้นปราศจากการพิสูจน์ ดังนั้น

$$P(X = b) = F(b) - F(b-)$$

ในเมื่อ $F(b-)$ เป็น limit ด้านซ้ายมือของ $F(x)$ ที่ $X = b$ นั่นคือ ความน่าจะเป็นที่ $X = b$ คือความสูงของขั้นที่ $F(x)$ อยู่ที่ $X = b$ ดังนั้น ถ้าฟังก์ชันการแจกแจง $F(x)$ ต่อเนื่องที่ $X = b$ แล้ว $P(X = b) = 0$

ง. $F(x)$ ต่อเนื่องไปทางขวาที่แต่ละจุด x

เพื่อที่จะพิสูจน์คุณสมบัตินี้ ให้ $h > 0$

$$\lim_{h \to 0} P(a < X \le a+h) = \lim_{h \to 0} F(a+h) - F(a) = F(\emptyset) = 0$$

เพราะว่าขณะ h เข้าใกล้ศูนย์ limit ของเซท $\{x ; a < X \le a + h\}$ เป็นเซทว่างเปล่า เพราะฉะนั้น เราเขียน

$$0 = F(a+) - F(a)$$

ในเมื่อ $F(a+)$ เป็น limit ทางด้านขวามือของ $F(x)$ ที่ $X = a$ ดังนั้น $F(x)$ ต่อเนื่องไปทางขวาที่ทุก ๆ จุด $X = a$

ตัวอย่างที่ 4 ให้ฟังก์ชันการแจกแจงกำหนดได้

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x+1}{2} & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

แล้ว

$$P(-3 < X \le 1/2) = F(1/2) - F(-3) = 3/4 - 0 = 3/4$$

และ $$P(X = 0) = F(0) - F(0-) = \frac{1}{2} - 0 = 1/2$$

ทฤษฎีที่ 2 (ก) ให้ F เป็น cdf ตัวแปรเชิงสุ่มชนิดต่อเนื่อง มี pdf f แล้ว

$$f(x) = \frac{d}{dx} F(x)$$

สำหรับ x ทั้งหมดที่อนุพันธ์ของ F

(ข) ให้ X เป็นตัวแปรเชิงสุ่มชนิดไม่ต่อเนื่อง พร้อมด้วยค่าที่เป็นไปได้ $x_1, x_2, \ldots$ และสมมติว่า $x_1 < x_2 < \ldots$ และ F เป็น cdf ของ X แล้ว

$$f(x_j) = P(X = x_j) = F(x_j) - F(x_{j-})$$

พิสูจน์ (ก) $F(x) = P(X \le x) = \int_{-\infty}^{x} f(s)\,ds$ ใช้ทฤษฎีพื้นฐานของแคลคูลัส เราได้ $F'(x) = f(x)$

(ข) เนื่องจากว่า เราสมมติ $x_1 < x_2 < \ldots$ เราได้

$$\begin{aligned} F(x_j) &= P(X = x_j \cup X = x_{j-1} \cup \ldots \cup X = x_1) \\ &= p(j) + p(j-1) + \ldots + p(1) \end{aligned}$$

และ

$$\begin{aligned} F(x_{j-1}) &= P(X = x_{j-1} \cup X = X_{j-2} \cup \ldots \cup X = x_1) \\ &= p(j-1) + p(j-2) + \ldots + p(1) \end{aligned}$$

ดังนั้น $F(x_j) - F(x_{j-1}) = P(X = x_j) = p(x_j) = f(x_j)$

ข้อสังเกต ให้เรากลับมาพิจารณาข้อ (ก) ของทฤษฎีข้างต้นอย่างย่อ ๆ กลับมายังนิยามของของอนุพันธ์ของฟังก์ชัน F

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h}$$

$$= \lim_{h \to 0^+} \frac{P(X \le x+h) - P(X \le x)}{h}$$

$$= \lim_{h \to 0^+} \frac{1}{h} [P(x < X \le x + h)]$$

ถ้า h มีค่าน้อยและเป็นบวก

$$F'(x) = f(x) \cong \frac{P(x < X \le x+h)}{h}$$

นั่นคือ f(x) เท่ากับ "จำนวนของความน่าจะเป็นในช่วง (x, x+h) ต่อความยาว h" เรียกว่าฟังก์ชันความหนาแน่นน่าจะเป็น

ตัวอย่างที่ 5 สมมติว่า ตัวแปรเชิงสุ่มชนิดต่อเนื่อง มี cdf F กำหนดได้

$$F(x) = 0 \qquad x \le 0$$
$$= 1 - e^{-x} \qquad x > 0$$

แล้ว $F'(x) = e^{-x}$ สำหรับ $x > 0$ ดังนั้น pdf f กำหนดได้โดย

$$f(x) = e^{-x}, \qquad x \ge 0$$
$$= 0 \quad \text{สำหรับค่าอื่น ๆ}$$

ตัวอย่างที่ 6 ให้ $f(x) = 1/2 \quad -1 < x < 1$ ศูนย์สำหรับค่าอื่น ๆ เป็น pdf ของตัวแปรเชิงสุ่ม X กำหนดให้ Y เป็นตัวแปรเชิงสุ่ม โดย $Y = X^2$ เราประสงค์ที่จะคำนวณหา pdf ของ Y ถ้า $y \ge 0$ ความน่าจะเป็น $P(Y \le y)$ equivalent กับ $P(X^2 \le y)$

$$= P(-\sqrt{y} \le X \le \sqrt{y})$$

ดังนั้น ฟังก์ชันการแจกแจงของ Y, $G(y) = P(Y \le y)$ กำหนดได้โดย

$$G(y) = 0 \qquad y < 0$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{2}\, dx = \sqrt{y} \qquad 0 \le y < 1$$

$$= 1 \qquad y \ge 1$$

เนื่องจากว่า Y เป็นตัวแปรเชิงสุ่มของชนิดต่อเนื่อง pdf ของ Y คือ $g(y) = G'(y)$ ที่จุดทั้งหมดของ continuity ของ g(y) ดังนั้น เราอาจเขียน

$$g(y) = \frac{1}{2\sqrt{y}}, \quad 0 < y < 1$$

$$= 0 \quad \text{สำหรับค่าอื่น ๆ}$$

ให้ตัวแปรเชิงสุ่ม X และ Y มี probability set function P(A) ในเมื่อ A เป็นเซทสองมิติ ถ้า A เป็น unbounded set $\{(u, v) ; u \le x, v \le y\}$ ในเมื่อ x และ y เป็นเลขจำนวนจริง เราได้

$$P(A) = P[(X, Y) \in A] = P(X \le x, Y \le y)$$

ฟังก์ชันนี้ของจุด (x, y) เรียกว่า ฟังก์ชันการแจกแจงของ X กับ Y และแสดงได้โดย

$$F(x, y) = P(X \le x, Y \le y)$$

ถ้า X และ Y เป็นตัวแปรเชิงสุ่มของชนิดต่อเนื่อง ซึ่งมี pdf f(x, y) แล้ว

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v)\, du\, dv$$

ดังนั้น ที่จุดของ continuity ของ f(x, y) เราได้

$$\frac{\partial^2 F(x, y)}{\partial x\, \partial y} = f(x, y)$$

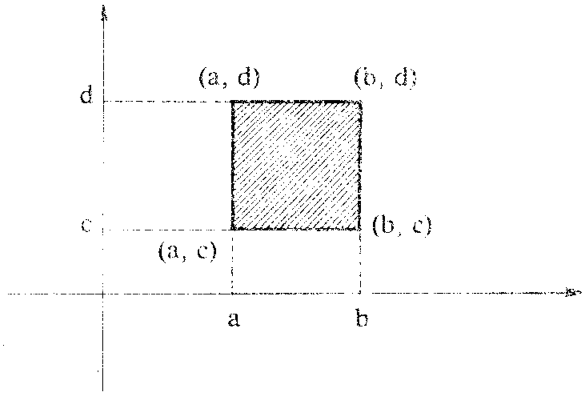
ในกรณี

$P(a < X \le b, c < Y \le d) = F(b, d) - F(b, c) - F(a, d) + F(a, c)$

สำหรับค่าคงที่ $a < b, c < d$ ที่เป็นเลขจำนวนจริงทั้งหมด

วิธีทำ ให้ $I = \{(x, y) : a < X \le b, c < Y \le d\}$

$$P\{(X, Y) \in I\} = P(X \le b, Y \le d) - P(X \le b, Y \le c)$$
$$- P(X \le a, Y \le d) + P(X \le a, Y \le c)$$
$$P\{(X, Y) \in I\} = F(b, d) - F(b, c) - F(a, d) + F(a, c)$$



กรณีฟังก์ชันการแจกแจงของ n ตัวแปรเชิงสุ่ม $X_1, X_2, \ldots, X_n$ เป็นฟังก์ชันเชิงจุด (point function)

$$F(x_1, x_2, \ldots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n)$$

ตัวอย่างที่ 7 ให้ $f(x, y, z) = e^{-(x+y+z)}$, $0 < x, y, z < \infty$, ศูนย์สำหรับค่าอื่น ๆ เป็น pdf ของตัวแปรเชิงสุ่ม X, Y และ Z แล้ว ฟังก์ชันการแจกแจงของ X, Y และ Z กำหนดโดย

$$\begin{aligned}
F(x, y, z) &= P(X \le x, Y \le y, Z \le z) \\
&= \int_0^z \int_0^y \int_0^x e^{-u-v-w}\, du\, dv\, dw \\
&= (1 - e^{-x})(1 - e^{-y})(1 - e^{-z}) \qquad 0 \le x, y, z < \infty \\
&= 0 \qquad \text{สำหรับค่าอื่น ๆ}
\end{aligned}$$

pdf ของ X, Y, Z คือ

$$\frac{\partial^3 F(x, y, z)}{\partial x \partial y \partial z} = f(x, y, z)$$

1.9 ตัวแบบความน่าจะเป็นสำหรับการทดลองสุ่ม

ในการทดลองหนึ่ง เลือกจุดหนึ่งโดยสุ่มที่จุด X จากช่วง[illegible]

เส้นจริง ๆ ดังนั้น sample space S คือ $\{x ; a \le x \le b\}$ และตัวแปรเชิงสุ่ม X เป็น identity function ที่ได้กำหนดบน S สมมุติว่า ช่วงระหว่าง A เป็นเซทย่อยของ S ความน่าจะเป็นของเหตุการณ์ A เป็นสัดส่วนกับความยาวของ A ดังนั้น ถ้า A เป็นช่วงระหว่าง $[a, x]$, $x \le b$ แล้ว

$$P(A) = P(X \in A) = P(a \le X \le x) = c(x-a)$$

ดังนั้น $c = 1/(b-a)$ และเป็นตัวแบบความน่าจะเป็นของการทดลองนี้ ถ้าเราเอาฟังก์ชันการแจกแจงของ X, $F(x) = P(X \le x)$ เป็น

$$\begin{aligned} F(x) &= 0 & x < a \\ &= \frac{x-a}{b-a} & a \le x \le b \\ &= 1 & x > b \end{aligned}$$

ดังนั้น pdf ของ X, $F(x) = F'(x)$ เขียนได้

$$\begin{aligned} f(x) &= \frac{1}{b-a}, \quad a \le x \le b \\ &= 0 \quad \text{สำหรับค่าอื่น ๆ} \end{aligned}$$

อนุพันธ์ของ $F(x)$ หาค่าไม่ได้ที่ $x = a$ หรือที่ $x = b$ แต่เซท $\{x ; x = a, b\}$ เป็นเซทของมาตราวัดความน่าจะเป็นศูนย์ และเลือกที่จะกำหนด $f(x)$ เท่ากับ $1/(b-a)$ ที่สองจุดเหล่านั้น สังเกตว่า pdf นี้คงที่บน S ถ้า pdf ของหนึ่งตัวแปร หรือมากกว่าของชนิดต่อเนื่องหรือไม่ต่อเนื่องเป็นค่าบน sample space S จึงกล่าวได้ว่า ความน่าจะเป็นเป็นการแจกแจง uniformly ทั้งหมดบน S เพราะฉะนั้น ในตัวอย่างข้างต้น X มีการแจกแจง uniform ตลอดช่วงระหว่าง $[a, b]$

ในกรณีตัวแปรเชิงสุ่ม X ชนิดไม่ต่อเนื่องที่มีการแจกแจงความน่าจะเป็น uniformly ตลอด k จุดใน sample space $S = \{x, x = 1, 2, \ldots, k\}$ pdf ของ X คือ $f(x) = \frac{1}{k}$, $x \in A$ ศูนย์สำหรับค่าอื่น ๆ pdf ชนิดนี้ใช้บรรยายตัวแบบความน่าจะเป็นสำหรับการทดลองเชิงสุ่มเมื่อแต่ละ k ผลลัพธ์มีความน่าจะเป็นเหมือนกันคือ $1/k$

ตัวแบบความน่าจะเป็นสามารถปรับให้เข้ากับภาวะทั่ว ๆ ไปได้มาก ให้ probability set function $P(A)$ ถูกกำหนดบน sample space S, S อาจเป็นเซทในหนึ่งหรือสองหรือหลาย ๆ มิติ ให้ S แบ่งออกเป็น k เซทย่อย $A_1, A_2, \ldots, A_k$ ที่ mutually exclusive และ exhausive สมมติว่าการทดลองเชิงสุ่มมีเหตุการณ์ที่แต่ละเหตุการณ์ A_i, $i = 1, 2, \ldots, k$ เป็น mutually exclusive และ exhausive มีความน่าจะเป็นเหมือนกัน $P(A) = 1/k$, $i = 1, 2, \ldots, k$ ให้เหตุการณ์ E เป็นผลรวมของ r เหตุการณ์ที่ mutually exclusive

$$E = A_1 \cup A_2 \cup \ldots \cup A_r \qquad r \leq k$$

แล้ว

$$P(E) = P(A_1) + P(A_2) + \ldots + P(A_r) = r/k$$

โดยทั่ว ๆ ไป จำนวนเต็ม k เรียกว่าจำนวนทั้งหมดของวิธีการซึ่งการทดลองเชิงสุ่มสามารถสิ้นสุดลง และจำนวนเต็ม r เรียกว่า จำนวนของวิธีการที่สอดคล้องกับเหตุการณ์ E ดังนั้น P(E) เท่ากับจำนวนของวิธีการที่สอดคล้องต่อเหตุการณ์ E หารด้วยจำนวนทั้งหมดของวิธีการในที่ซึ่งการทดลองสามารถสิ้นสุดลง เพื่อกำหนดความน่าจะเป็น r/k ต่อเหตุการณ์ E เราต้องสมมติว่า แต่ละเหตุการณ์ $A_1, A_2, \ldots, A_k$ ที่ mutually exclusive และ exhausive กัน มีความน่าจะเป็น 1/k และเป็นเงื่อนไขของตัวแบบความน่าจะเป็น

ตัวอย่างที่ 1 เลือกไพ่ใบหนึ่งจากไพ่สำรับหนึ่ง 52 ใบ โดยวิธีสุ่ม sample space คือผลรวมของ $k = 52$ และสมมติว่าแต่ละผลลัพธ์มีความน่าจะเป็น $\frac{1}{52}$ ถ้า E_1 เป็นเซทของผลลัพธ์ที่เป็นโพดำ $P(E_1) = \frac{13}{52} = \frac{1}{4}$ เพราะว่ามี $r_1 = 13$ ใบในสำรับ ถ้า E_2 เป็นเซทของผลลัพธ์ที่เป็นคิง $P(E_2) = \frac{4}{52} = \frac{1}{13}$ เพราะว่ามี $r_2 = 4$ ในสำรับ นั่นคือ $\frac{1}{13}$ คือ ความน่าจะเป็นของการดึงไพ่ใบหนึ่งเป็นคิง สมมติดึงไพ่ห้าใบโดยสุ่มและปราศจากการใส่คืนจากไพ่สำรับนี้ เราต้องคิดเสียว่าไพ่แต่ละห้าใบเป็นผลลัพธ์หนึ่งใน sample space และแต่ละผลลัพธ์ของผลลัพธ์เหล่านี้มีความน่าจะเป็นเท่า ๆ กัน ถ้า E_1 เป็นเซทของผลลัพธ์ในที่ซึ่งแต่ละใบที่ได้เป็นโพดำ $P(E_1)$ เท่ากับจำนวน r_1 ของโพดำทั้งหมดที่ได้ หารด้วยจำนวนทั้งหมด k ของไพ่ห้าใบที่ได้

$$r_1 = \binom{13}{5} = \frac{13!}{5!8!} \qquad k = \binom{52}{5} = \frac{52!}{5!47!}$$

โดยทั่ว ๆ ไป ถ้า n เป็นเลขจำนวนเต็มบวกและถ้า x เป็นเลขจำนวนเต็มบวกพร้อมด้วย $x \leq n$ แล้ว สัมประสิทธิ์ทวินาม

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

เท่ากับจำนวนของหมู่ของ n สิ่งเลือก x สิ่งในครั้งหนึ่ง
ดังนั้น

$$P(E_1) = \frac{\binom{13}{5}}{\binom{52}{5}} = \frac{\frac{13!}{5!8!}}{\frac{52!}{5!\,47!}}$$

$$= \frac{(13)(12)(11)(10)(9)}{(52)(51)(50)(49)(48)} = 0.0005$$

ต่อไปให้ E_2 เป็นเซทของผลลัพธ์ในที่ซึ่งอย่างน้อยไพ่หนึ่งใบเป็นโพดำแล้ว $\bar{E}_2$ เป็นเซทของผลลัพธ์ในที่ซึ่งไม่มีไพ่เป็นโพดำ มี $\bar{r}_2 = \binom{39}{5}$ ผลลัพธ์

$$P(\bar{E}_2) = \frac{\binom{39}{5}}{\binom{52}{5}} \text{ และ } P(E_2) = 1 - P(\bar{E}_2)$$

สมมติ E_3 เป็นเซทของผลลัพธ์ในที่ซึ่งไพ่สามใบเป็นคิงและสองใบเป็นควีน เราสามารถเลือกสามคิงในหนึ่งวิธีใดวิธีหนึ่งของ $\binom{4}{3}$ วิธี และสองควีนในหนึ่งวิธีใดวิธีหนึ่งของ $\binom{4}{2}$ วิธี โดยหลักของการนับ จำนวนของผลลัพธ์ใน E_3 คือ $r_3 = \binom{4}{3}\binom{4}{2}$ ดังนั้น $P(E_3) = \binom{4}{3}\binom{4}{2} / \binom{52}{5}$ ให้ E_4 เป็นเซทของผลลัพธ์ในที่ซึ่งมีสองคิง, สองควีนและหนึ่งแจ็คแล้ว

$$P(E_4) = \frac{\binom{4}{2}\binom{4}{2}\binom{4}{1}}{\binom{52}{5}}$$

ตัวอย่างที่ 2 กล่องใบหนึ่งมีหลอดไฟอยู่ 100 หลอด เลือกหลอดไฟมาห้าหลอดโดยสุ่มและทดสอบดู ถ้าหากว่าหลอดไฟทั้งห้าหลอดดีทั้งหมดเป็นอันยอมรับกล่องใบนั้น ถ้าหากว่าในกล่องใบนั้นมีหลอดไฟเสีย 20 หลอด ความน่าจะเป็นของการยอมรับกล่องนั้นภายใต้เงื่อนไขที่เหมาะสมคือ

$$\frac{\binom{80}{5}}{\binom{100}{5}} = 0.32$$

โดยทั่ว ๆ ไป ให้ X เป็นจำนวนของหลอดไฟเสียระหว่างที่ทำการทดสอบ 5 หลอด space ของ X คือ $S = \{ x ; x = 0,1,2,3,4,5 \}$ และ pdf ของ X กำหนดให้ได้โดย

$$f(x) = P(X = x) = \frac{\binom{20}{x}\binom{80}{5-x}}{\binom{100}{5}}, \; x = 0,1,2,3,4,5$$

$$= 0 \quad \text{สำหรับค่าอื่น ๆ}$$

นี่เป็นตัวอย่างของชนิดไม่ต่อเนื่องของการแจกแจงซึ่งเรียกว่า hyper geometric distribution

ตัวอย่างที่ 3 ให้ X เป็นตัวแปรเชิงสุ่มซึ่งมีการแจกแจง uniformly บนช่วง -1 ถึง 1 จงหา pdf ของฟังก์ชันกำหนดได้เป็น

(ก) $Y = -\log_e |X|$ (ข) $Y = \cos(\pi X)$

วิธีทำ (ก) pdf ของตัวแปรเชิงสุ่ม X มีการแจกแจง uniformly บนช่วง (-1, 1) เป็น

$$f(x) = \frac{1}{2}, \; -1 < x < 1$$

$= 0$ สำหรับค่าอื่น ๆ

ในที่นี้ $Y = \log_e |X|$

เราได้ $G(Y) = P(Y \le y) = P(-\log_e |X| \le y)$

$= P(|X| > e^{-y})$

$= 1 - P(|X| \le e^{-y})$

$= 1 - P(-e^{-y} \le X \le e^{-y})$

$= 1 - [F(e^{-y}) - F(-e^{-y})]$

$g(y) = G'(y) = e^{-y} f(e^{-y}) + e^{-y} f(-e^{-y})$

$= e^{-y} [f(-e^{-y}) + f(e^{-y})]$

$= e^{-y} \left[\frac{1}{2} + \frac{1}{2}\right]$

$= e^{-y}$ สำหรับ $y > 0$

$= 0$ สำหรับค่าอื่น ๆ

(ข)

$G(y) = P(Y \le y) = P(\text{Cos}\, \pi X \le y)$

$= P(X \le \frac{1}{\pi} \text{Cos}^{-1} y)$

$= F(\frac{1}{\pi} \text{Cos}^{-1} y)$

$g(y) = -\frac{1}{\pi \sqrt{1 - y^2}} f\left(\frac{1}{\pi} \text{Cos}^{-1} y\right)$

$= -\frac{1}{2\pi \sqrt{1 - y^2}} \quad -1 < y < 1$

$= 0$ สำหรับค่าอื่น ๆ