เอกสารประกอบการสอนรายวิชา
103103 ความน่าจะเป็นและสถิติ
(Probability and Statistics)

อาจารย์ ดร. ธิดารัตน์ อารีรักษ์
สาขาวิชาคณิตศาสตร์ สำนักวิชาวิทยาศาสตร์
มหาวิทยาลัยเทคโนโลยีสุรนารี

นิยามต่อไปจะกล่าวถึง ปริภูมิตัวอย่าง จุดตัวอย่าง เหตุการณ์ และเหตุการณ์ที่ไม่เกิดร่วมกัน

<u>นิยามที่ 1.1.2</u>

ปริภูมิตัวอย่าง (sample space) หมายถึง เซตของผลลัพธ์ที่เป็นไปได้ทั้งหมดที่เกิดจากการทดลองสุ่ม เขียนแทนด้วยสัญลักษณ์ S

จุดตัวอย่าง (sample point) หมายถึง สมาชิกแต่ละตัวของปริภูมิตัวอย่าง

เหตุการณ์ (event) หมายถึง สับเซตหรือเซตย่อยของปริภูมิตัวอย่าง นิยมเขียนแทนด้วยตัวอักษรภาษาอังกฤษตัวใหญ่ เช่น A, B หรือ C เป็นต้น

เหตุการณ์ที่ไม่เกิดร่วมกัน (mutually disjoint event) เหตุการณ์ $A_1, \ldots, A_n$ จะเรียกว่าเป็นเหตุการณ์ที่ไม่เกิดร่วมกัน ถ้า $A_i \cap A_j = \varnothing, \forall i \neq j$ โดยที่ $i, j = 1, 2, \ldots, n$

<u>หมายเหตุ</u>

เหตุการณ์ที่เป็นไปไม่ได้ แทนได้ด้วยเซตว่างหรือสัญลักษณ์ $\varnothing$

<u>ตัวอย่างที่ 1.1.2</u>

(1) พิจารณาการโยนลูกเต๋า 1 ลูก 1 ครั้ง สังเกตแต้มที่ลูกเต๋าออก จงเขียนปริภูมิตัวอย่างและเหตุการณ์ที่ได้แต้มมากกว่า 3

ปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

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

ถ้าให้ E แทนเหตุการณ์ที่ลูกเต๋าได้แต้มมากกว่า 3 ดังนั้น

$$E = \{4, 5, 6\}$$

(2) พิจารณาการมีบุตร 3 คน ของสามีภรรยาคู่หนึ่ง สังเกตเพศของบุตร จงเขียนปริภูมิตัวอย่างและเหตุการณ์ที่ลูกเป็นเพศชายอย่างน้อย 2 คน

ปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

$S = \{$ ชชช, ชชญ, ชญช, ชญญ, ญชช, ญชญ, ญญช, ญญญ $\}$

ถ้าให้ E แทนเหตุการณ์ที่ลูกเป็นเพศชายอย่างน้อย 2 คน ดังนั้น

$E = \{$ ชชช, ชชญ, ชญช, ญชช $\}$

(3) พิจารณาการมีบุตร 3 คน ของสามีภรรยาคู่หนึ่ง สังเกตจำนวนบุตรชาย จงเขียนปริภูมิตัวอย่างและเหตุการณ์ที่ลูกเป็นเพศชายอย่างน้อย 2 คน

ปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

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

ถ้าให้ E แทนเหตุการณ์ที่ลูกเป็นเพศชายอย่างน้อย 2 คน ดังนั้น

$$E = \{2, 3\}$$

(4) พิจารณาการปล่อยอนุภาคของสารกัมมันตรังสีในเวลา 1 นาที สังเกตจำนวนอนุภาคของสารกัมมันตรังสีที่แผ่ออกมา จงเขียนปริภูมิตัวอย่างและเหตุการณ์ที่มีจำนวนอนุภาคน้อยกว่า 10 อนุภาค

ปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

$$S=\{0,1,2,\ldots\}$$

ถ้าให้ E แทนเหตุการณ์ที่มีจำนวนอนุภาคน้อยกว่า 10 อนุภาค ดังนั้น

$$E=\{0,1,2,3,4,5,6,7,8,9\}$$

■

<u>ข้อสังเกต</u> จากตัวอย่างที่ 1.1.2 จะเห็นว่า

1. จำนวนสมาชิกในปริภูมิตัวอย่างเป็นไปได้ทั้งจำนวนจำกัด และจำนวนอนันต์
2. ปริภูมิตัวอย่างของการทดลองสุ่มเดียวกันสามารถเขียนได้หลายแบบขึ้นอยู่กับสิ่งที่สนใจหรือสิ่งที่สังเกต

1.2 หลักการนับพื้นฐาน

การนับจำนวนสมาชิกของเหตุการณ์ หรือจำนวนสมาชิกของปริภูมิตัวอย่างเป็นพื้นฐานที่สำคัญสำหรับการคำนวณหาค่าความน่าจะเป็น ในบางเหตุการณ์เราสามารถแจกแจงสมาชิกออกมาได้ทั้งหมด ทำให้นับจำนวนสมาชิกของเหตุการณ์ได้โดยตรง แต่ในบางเหตุการณ์ที่มีจำนวนสมาชิกมาก ๆ การแจกแจงสมาชิกออกมาทั้งหมดจึงเป็นเรื่องที่ยุ่งยาก ดังนั้น ในหัวข้อนี้จะกล่าวถึงหลักการนับพื้นฐานที่สำคัญซึ่งประกอบไปด้วย หลักการบวก หลักการคูณ การเรียงสับเปลี่ยน และการจัดหมู่ ซึ่งช่วยให้คำนวณหาจำนวนสมาชิกในเหตุการณ์ได้โดยที่ไม่ต้องแจกแจงสมาชิกของเหตุการณ์ออกมาทั้งหมด

1.2.1 หลักการบวก

หลักการบวกเป็นหลักการนับพื้นฐานที่สำคัญอย่างหนึ่ง ซึ่งใช้สำหรับนับจำนวนสมาชิกของเหตุการณ์ที่ไม่เกิดร่วมกัน

หลักการบวก (addition principle)

กำหนดให้ E เป็นเหตุการณ์ใด ๆ ที่สามารถแบ่งออกได้เป็น k กรณี ที่ไม่เกิดร่วมกัน โดยที่

กรณีที่ 1 สามารถเกิดได้ n_1 วิธี

กรณีที่ 2 สามารถเกิดได้ n_2 วิธี

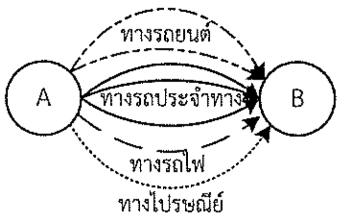
$\vdots$

กรณีที่ k สามารถเกิดได้ n_k วิธี

ดังนั้น เหตุการณ์ E สามารถเกิดขึ้นได้ทั้งหมด $n_1+n_2+\ldots+n_k$ วิธี

ตัวอย่างที่ 1.2.1 การขนส่งสินค้าชนิดหนึ่งจากโรงงานผู้ผลิตไปยังสำนักงานขายสามารถขนส่งได้ทั้งหมด 4 วิธี คือ ทางรถยนต์ ทางรถประจำทาง ทางรถไฟ และทางไปรษณีย์ โดยทางรถยนต์สามารถเลือกบริษัทขนส่งได้ 2 บริษัท ทางรถประจำทางสามารถเลือกบริษัทขนส่งได้ 3 บริษัท ส่วนทางรถไฟและทางไปรษณีย์มีบริษัทขนส่งเพียงอย่างละ 1 บริษัทเท่านั้น จงหาจำนวนวิธีทั้งหมดในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย

วิธีทำ ทางเลือกในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย แสดงได้ในรูปที่ 1.2.1



รูปที่ 1.2.1 ทางเลือกในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย

การขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย สามารถแบ่งออกได้เป็น 4 กรณี ที่ไม่เกิดร่วมกันดังนี้

กรณีที่ 1: ถ้าเลือกการขนส่งทางรถยนต์สามารถขนส่งได้ทั้งหมด 2 วิธี

กรณีที่ 2: ถ้าเลือกการขนส่งทางรถประจำทางสามารถขนส่งได้ทั้งหมด 3 วิธี

กรณีที่ 3: ถ้าเลือกการขนส่งทางรถไฟสามารถขนส่งได้ทั้งหมด 1 วิธี

กรณีที่ 4: ถ้าเลือกการขนส่งทางไปรษณีย์สามารถขนส่งได้ทั้งหมด 1 วิธี

จากหลักการบวก สรุปได้ว่าจำนวนวิธีในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขายมีทั้งหมด $2+3+1+1=7$ วิธี ■

ตัวอย่างที่ 1.2.2 พิจารณาการทอดลูกเต๋า 2 ลูกที่แตกต่างกัน จงหาจำนวนวิธีที่ลูกเต๋าจะขึ้นแต้ม 1 อย่างน้อย 1 ลูก

วิธีทำ เหตุการณ์ที่ลูกเต๋าจะขึ้นแต้ม 1 อย่างน้อย 1 ลูกแบ่งออกได้ 2 กรณี ที่ไม่เกิดร่วมกันดังนี้

กรณีที่ 1: ลูกเต๋าขึ้นแต้ม 1 จำนวน 1 ลูก เกิดได้ทั้งหมด 10 วิธี

กรณีที่ 2: ลูกเต๋าขึ้นแต้ม 1 จำนวน 2 ลูก เกิดได้ทั้งหมด 1 วิธี

จากหลักการบวก สรุปได้ว่าจำนวนวิธีที่ลูกเต๋า 2 ลูกที่แตกต่างกันจะขึ้นแต้ม 1 อย่างน้อย 1 ลูก เกิดได้ทั้งหมด $10+1=11$ วิธี ■

1.2.2 หลักการคูณ

หลักการคูณเป็นหลักการนับพื้นฐานที่สำคัญอีกอย่างหนึ่ง ซึ่งใช้สำหรับนับจำนวนสมาชิกของเหตุการณ์ที่สามารถแบ่งเป็นขั้นตอนได้

หลักการคูณ (multiplication principle)

กำหนดให้ E เป็นเหตุการณ์ใด ๆ ที่สามารถแบ่งออกเป็นขั้นตอนได้ k ขั้นตอน โดยที่

ขั้นตอนที่ 1 สามารถเกิดได้ n_1 วิธี

ขั้นตอนที่ 2 สามารถเกิดได้ n_2 วิธี

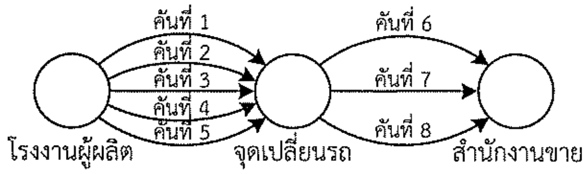
$\vdots$

ขั้นตอนที่ k สามารถเกิดได้ n_k วิธี

ดังนั้น เหตุการณ์ E สามารถเกิดขึ้นได้ทั้งหมด $n_1 \times n_2 \times \ldots \times n_k$ วิธี

ตัวอย่างที่ 1.2.3 การขนส่งสินค้าชนิดหนึ่งด้วยรถยนต์จากโรงงานผู้ผลิตไปยังสำนักงานขายจะต้องมีการหยุดพักเพื่อเปลี่ยนรถ ณ จุดเปลี่ยนรถ โดยที่รถยนต์ที่ใช้ในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังจุดเปลี่ยนรถมีทั้งหมด 5 คัน และรถยนต์ที่ใช้ในการขนส่งสินค้าจากจุดเปลี่ยนรถไปยังสำนักงานขายมีทั้งหมด 3 คัน จงหาจำนวนวิธีทั้งหมดในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย

วิธีทำ ทางเลือกในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย โดยมีการหยุดพักเพื่อเปลี่ยนรถ แสดงได้ในรูปที่ 1.2.2



รูปที่ 1.2.2 ทางเลือกในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย โดยมีการหยุดพักเพื่อเปลี่ยนรถ

การขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย โดยมีการหยุดพักเพื่อเปลี่ยนรถ สามารถแบ่งออกได้เป็น 2 ขั้นตอนดังนี้

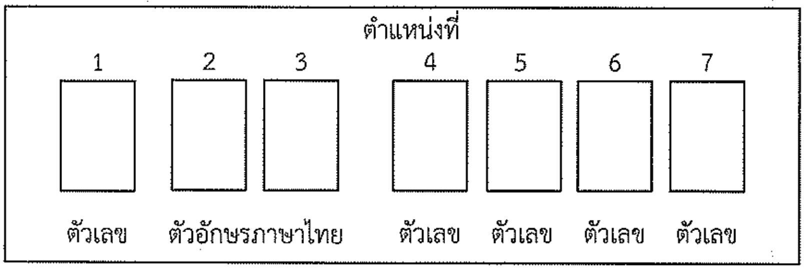
ขั้นตอนที่ 1: เดินทางจากโรงงานผู้ผลิตไปจุดเปลี่ยนรถสามารถเดินทางได้ทั้งหมด 5 วิธี

ขั้นตอนที่ 2: เดินทางจากจุดเปลี่ยนรถไปสำนักงานขายสามารถเดินทางได้ทั้งหมด 3 วิธี

จากหลักการคูณ สรุปได้ว่าจำนวนวิธีในการขนส่งสินค้าจากโรงงานผู้ผลิตไปยังสำนักงานขาย โดยมีการหยุดพักเพื่อเปลี่ยนรถมีทั้งหมด $5 \times 3 = 15$ วิธี ■

<u>**ตัวอย่างที่ 1.2.4**</u> การพิมพ์ป้ายทะเบียนรถยนต์ประกอบด้วยตัวเลข 1 ตัว ตามด้วยตัวอักษรภาษาไทยที่แตกต่างกัน 2 ตัว ตามด้วยตัวเลขอีก 4 ตัว โดยกำหนดให้ตัวเลขในตำแหน่งแรกและตำแหน่งที่ 4 ไม่เท่ากับ 0 ดังรูปที่ 1.2.3 จงหาจำนวนวิธีทั้งหมดในการพิมพ์ป้ายทะเบียนรถยนต์

วิธีทำ รูปป้ายทะเบียนรถยนต์ แสดงได้ในรูปที่ 1.2.3



รูปที่ 1.2.3 ป้ายทะเบียนรถยนต์

การพิมพ์ป้ายทะเบียนรถยนต์สามารถแบ่งออกได้เป็น 7 ขั้นตอนดังนี้

ขั้นตอนที่ 1: พิมพ์ตัวเลขลงไปในตำแหน่งที่ 1 ทำได้ทั้งหมด 9 วิธี (เนื่องจากตัวเลขในตำแหน่งนี้ต้องไม่เท่ากับ 0 ดังนั้นตัวเลขที่สามารถพิมพ์ในตำแหน่งนี้ได้จึงมีเพียง 1, 2, 3, 4, 5, 6, 7, 8 และ 9)

ขั้นตอนที่ 2: พิมพ์ตัวอักษรภาษาไทยลงไปในตำแหน่งที่ 2 ทำได้ทั้งหมด 44 วิธี (เนื่องจากตัวอักษรภาษาไทยมีทั้งหมด 44 ตัว)

ขั้นตอนที่ 3: พิมพ์ตัวอักษรภาษาไทยลงไปในตำแหน่งที่ 3 ทำได้ทั้งหมด 43 วิธี (เนื่องจากต้องการให้ตัวอักษรภาษาไทยทั้งสองตัวแตกต่างกัน แต่เราได้พิมพ์ตัวอักษรภาษาไทยลงไปในตำแหน่งที่ 2 แล้ว 1 ตัว ดังนั้นในตำแหน่งที่ 3 จึงเหลือตัวอักษรภาษาไทยที่ใช้ได้อีกเพียง 43 ตัว เท่านั้น)

ขั้นตอนที่ 4: พิมพ์ตัวเลขลงไปในตำแหน่งที่ 4 ทำได้ทั้งหมด 9 วิธี (เนื่องจากตัวเลขในตำแหน่งนี้ต้องไม่เท่ากับ 0 ดังนั้นตัวเลขที่สามารถพิมพ์ในตำแหน่งนี้ได้จึงมีเพียง 1, 2, 3, 4, 5, 6, 7, 8 และ 9)

ขั้นตอนที่ 5: พิมพ์ตัวเลขลงไปในตำแหน่งที่ 5 ทำได้ทั้งหมด 10 วิธี (ตัวเลขที่สามารถใส่ในตำแหน่งนี้ได้ประกอบไปด้วย 0, 1, 2, 3, 4, 5, 6, 7, 8 และ 9)

ขั้นตอนที่ 6: พิมพ์ตัวเลขลงไปในตำแหน่งที่ 6 ทำได้ทั้งหมด 10 วิธี (ใช้วิธีการคิดเช่นเดียวกับขั้นตอนที่ 5)

ขั้นตอนที่ 7: พิมพ์ตัวเลขลงไปในตำแหน่งที่ 7 ทำได้ทั้งหมด 10 วิธี (ใช้วิธีการคิดเช่นเดียวกับขั้นตอนที่ 5 และ 6)

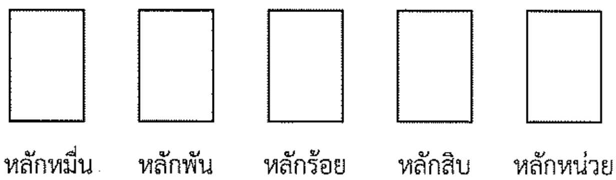
จากหลักการคูณ สรุปได้ว่าจำนวนวิธีในการพิมพ์ป้ายทะเบียนรถยนต์สามารถทำได้ทั้งหมด $9\times 44\times 43\times 9\times 10\times 10\times 10=153,252,000$ วิธี ■

<u>ข้อสังเกต</u>

การหาจำนวนสมาชิกของเหตุการณ์สำหรับตัวอย่างที่ 1.2.1 และ 1.2.2 ใช้หลักการบวกเพียงอย่างเดียว ส่วนตัวอย่างที่ 1.2.3 และ 1.2.4 ใช้หลักการคูณเพียงอย่างเดียว แต่เหตุการณ์บางเหตุการณ์มีความซับซ้อน ไม่สามารถใช้หลักการบวก หรือหลักการคูณเพียงอย่างใดอย่างหนึ่งในการนับจำนวนสมาชิกได้ เช่น ตัวอย่างที่ 1.2.5 จะเป็นตัวอย่างที่ต้องใช้ทั้งหลักการบวก และหลักการคูณในการหาจำนวนสมาชิกของเหตุการณ์

<u>ตัวอย่างที่ 1.2.5</u> จงหาจำนวนเต็มคู่ที่มีค่าตั้งแต่ 20,000 ถึง 70,000 และไม่มีตัวเลขในหลักใดซ้ำกันเลย

วิธีทำ จำนวนที่ประกอบด้วยตัวเลข 5 หลัก แสดงได้ในรูปที่ 1.2.4



รูปที่ 1.2.4 จำนวนที่ประกอบด้วยตัวเลข 5 หลัก

จำนวนเต็มคู่ที่มีค่าตั้งแต่ 20,000 ถึง 70,000 และไม่มีตัวเลขในหลักใดซ้ำกันเลย สามารถแบ่งออกได้เป็น 2 กรณีที่ไม่เกิดร่วมกันดังนี้

<u>กรณีที่ 1:</u> หลักหมื่นเป็นตัวเลข 2, 4 หรือ 6 เนื่องจากจำนวนที่เราต้องการเป็นจำนวนที่มี 5 หลัก ดังนั้น เราสามารถแบ่งออกได้เป็น 5 ขั้นตอนดังนี้

ขั้นตอนที่ 1: พิจารณาหลักหมื่น มีตัวเลขที่เป็นไปได้ทั้งหมด 3 ตัว คือ 2, 4 หรือ 6

ขั้นตอนที่ 2: พิจารณาหลักหน่วย เนื่องจากเราต้องการจำนวนเต็มคู่ (ตัวเลขที่เป็นไปได้ในหลักนี้มีทั้งหมด 5 ตัว คือ 0, 2, 4, 6 หรือ 8) และไม่มีตัวเลขในหลักใดซ้ำกัน แต่เราเลือกตัวเลขในหลักหมื่นเป็น 2, 4 หรือ 6 ไปแล้ว ดังนั้น ในหลักนี้จึงมีตัวเลขที่เป็นไปได้เพียง 4 ตัวเท่านั้น

ขั้นตอนที่ 3: พิจารณาหลักพัน เนื่องจากเราต้องการจำนวนซึ่งไม่มีตัวเลขในหลักใดซ้ำกัน แต่เราเลือกตัวเลขในหลักหมื่น และหลักหน่วยไปแล้ว 2 ตัว ดังนั้น ในหลักนี้จึงมีตัวเลขที่เป็นไปได้เพียง 8 ตัวเท่านั้น

ขั้นตอนที่ 4: พิจารณาหลักร้อย ในทำนองเดียวกับขั้นตอนที่ 3 หลักนี้จึงมีตัวเลขที่เป็นไปได้ทั้งหมด 7 ตัว

ขั้นตอนที่ 5: พิจารณาหลักสิบ ในทำนองเดียวกับขั้นตอนที่ 3 และ 4 หลักนี้จึงมีตัวเลขที่เป็นไปได้ทั้งหมด 6 ตัว

จากหลักการคูณ สรุปได้ว่าจำนวนที่สอดคล้องกับกรณีที่ 1 มีทั้งหมด $3\times 4\times 8\times 7\times 6=4,032$ จำนวน

<u>กรณีที่ 2:</u> หลักหมื่นเป็นตัวเลข 3 หรือ 5 ทำนองเดียวกับกรณีที่ 1 เราสามารถแบ่งออกได้เป็น 5 ขั้นตอนดังนี้

ขั้นตอนที่ 1: พิจารณาหลักหมื่น มีตัวเลขที่เป็นไปได้ทั้งหมด 2 ตัว คือ 3 หรือ 5

ขั้นตอนที่ 2: พิจารณาหลักหน่วย เนื่องจากเราต้องการจำนวนเต็มคู่ (ตัวเลขที่เป็นไปได้ในหลักนี้มีทั้งหมด 5 ตัว คือ 0, 2, 4, 6 หรือ 8) และตัวเลขต้องไม่ซ้ำกับตัวเลขในหลักหมื่น เนื่องจากในหลักหมื่นตัวเลขที่เป็นไปได้คือ 3 หรือ 5 ดังนั้น ไม่ว่าในหลักหมื่นจะเป็นตัวเลขใดใน 2 ตัวนี้ก็ตาม หลักหน่วยจะมีตัวเลขที่เป็นไปได้ทั้งหมด 5 ตัวเช่นเดิม

ขั้นตอนที่ 3: พิจารณาหลักพัน เนื่องจากเราต้องการจำนวนซึ่งไม่มีตัวเลขในหลักใดซ้ำกัน แต่เราเลือกตัวเลขในหลักหมื่น และหลักหน่วยไปแล้ว 2 ตัว ดังนั้นในหลักนี้จึงมีตัวเลขที่เป็นไปได้เพียง 8 ตัวเท่านั้น

ขั้นตอนที่ 4: พิจารณาหลักร้อย ในทำนองเดียวกับขั้นตอนที่ 3 หลักนี้จึงมีตัวเลขที่เป็นไปได้ทั้งหมด 7 ตัว

ขั้นตอนที่ 5: พิจารณาหลักสิบ ในทำนองเดียวกับขั้นตอนที่ 3 และ 4 หลักนี้จึงมีตัวเลขที่เป็นไปได้ทั้งหมด 6 ตัว

จากหลักการคูณ สรุปได้ว่าจำนวนที่สอดคล้องกับกรณีที่ 2 มีทั้งหมด $2\times5\times8\times7\times6=3{,}360$ จำนวน

จากหลักการบวก สรุปได้ว่าจำนวนเต็มคู่ที่มีค่าตั้งแต่ 20,000 ถึง 70,000 และไม่มีตัวเลขในหลักใดซ้ำกันเลยมีทั้งหมด = จำนวนที่สอดคล้องกับกรณีที่ 1 + จำนวนที่สอดคล้องกับกรณีที่ 2 = $4{,}032+3{,}360=7{,}392$ จำนวน ■

<u>ข้อสังเกต</u> จากตัวอย่างที่ 1.2.5

1. เราไม่จำเป็นต้องพิจารณากรณีที่หลักหมื่นเป็นเลข 7 เนื่องจากจำนวนที่หลักหมื่นเป็นเลข 7 มีเพียง 70,000 จำนวนตัวเดียวเท่านั้น และ 70,000 มีตัวเลข 0 ซ้ำกันอยู่ จึงไม่สอดคล้องกับสิ่งที่โจทย์ต้องการ

2. การแบ่งเหตุการณ์ออกเป็นกรณีที่ไม่เกิดร่วมกันนั้น ยังสามารถพิจารณาโดยใช้หลักหน่วยเป็นเกณฑ์ได้เช่นเดียวกัน นั่นคือ สามารถแบ่งออกเป็น กรณีที่ 1 หลักหน่วยเป็นตัวเลข 0 หรือ 8 และกรณีที่ 2 หลักหน่วยเป็นตัวเลข 2, 4 หรือ 6

1.2.3 การเรียงสับเปลี่ยน

เหตุการณ์ที่เกี่ยวข้องกับการจัดเรียงลำดับ สามารถใช้การเรียงสับเปลี่ยนมาช่วยในการนับจำนวนสมาชิกของเหตุการณ์ได้ ซึ่งจะทำให้การนับจำนวนสมาชิกของเหตุการณ์ลดความยุ่งยากซับซ้อนลง เมื่อเทียบกับการใช้หลักการคูณในการนับจำนวนสมาชิกของเหตุการณ์

นิยามที่ 1.2.1 ***การเรียงสับเปลี่ยน (permutation)*** หมายถึง การนำสิ่งของบางส่วนหรือทั้งหมดมาจัดเรียงลำดับ โดยถือว่า "ลำดับมีความสำคัญ" นั่นคือ ถ้ามีการสลับลำดับหรือตำแหน่งจะทำให้เกิดการเรียงสับเปลี่ยนแบบใหม่ที่แตกต่างจากเดิม

ตัวอย่างที่ 1.2.6 ถ้ามีตัวอักษรภาษาอังกฤษทั้งหมด 3 ตัว คือ A, B และ C จะสามารถทำการเรียงสับเปลี่ยนได้ทั้งหมดกี่วิธี อะไรบ้างเมื่อ

(1) ต้องการเรียงสับเปลี่ยนตัวอักษร 1 ตัว

(2) ต้องการเรียงสับเปลี่ยนตัวอักษร 2 ตัว

(3) ต้องการเรียงสับเปลี่ยนตัวอักษร 3 ตัว

วิธีทำ (1) การเรียงสับเปลี่ยนตัวอักษร 1 ตัว จากตัวอักษรทั้งหมด 3 ตัว (A, B และ C) ทำได้ทั้งหมด 3 วิธี ดังนี้ A, B และ C

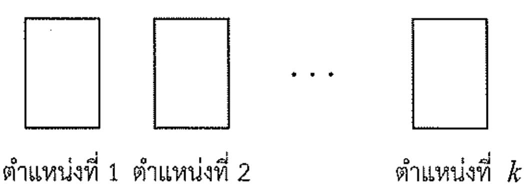
(2) การเรียงสับเปลี่ยนตัวอักษร 2 ตัว จากตัวอักษรทั้งหมด 3 ตัว (A, B และ C) ทำได้ทั้งหมด 6 วิธี ดังนี้ AB, AC, BA, BC, CA และ CB

(3) การเรียงสับเปลี่ยนตัวอักษร 3 ตัว จากตัวอักษรทั้งหมด 3 ตัว (A, B และ C) ทำได้ทั้งหมด 6 วิธี ดังนี้ ABC, ACB, BAC, BCA, CAB และ CBA ■

ข้อสังเกต

จากตัวอย่างที่ 1.2.6 จะสังเกตได้ว่าการเรียงสับเปลี่ยนตัวอักษร 3 ตัว จากตัวอักษรทั้งหมด 3 ตัว ABC และ ACB ถึงแม้จะมีตัวอักษรที่เหมือนกันทั้ง 3 ตัว แต่ลำดับของตัวอักษรแตกต่างกัน ทำให้การเรียงสับเปลี่ยนทั้ง 2 วิธีนี้แตกต่างกัน นั่นคือ สำหรับการเรียงสับเปลี่ยนลำดับมีความสำคัญ ถ้าลำดับแตกต่างกันจะทำให้เกิดการเรียงสับเปลี่ยนที่แตกต่างกันด้วย

พิจารณาการเรียงสับเปลี่ยนสิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง (ดังนั้น $k \le n$) แสดงได้ในรูปที่ 1.2.5



รูปที่ 1.2.5 ตำแหน่งของการเรียงสับเปลี่ยนสิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง

สามารถแบ่งออกได้เป็น k ขั้นตอน ดังนี้

ขั้นตอนที่ 1: ตำแหน่งที่ 1 สามารถจัดสิ่งของได้ทั้งหมด n วิธี

ขั้นตอนที่ 2: ตำแหน่งที่ 2 สามารถจัดสิ่งของได้ทั้งหมด $n-1$ วิธี

$$\vdots$$

ขั้นตอนที่ k: ตำแหน่งที่ k สามารถจัดสิ่งของได้ทั้งหมด $n-(k-1)$ วิธี

จากหลักการคูณ สรุปได้ว่าจำนวนวิธีการเรียงสับเปลี่ยนสิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง ทำได้ทั้งหมด $n(n-1)\cdots(n-k+1)$ วิธี

กำหนดให้ ${}^nP_k$ แทนจำนวนวิธีการเรียงสับเปลี่ยนสิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง จะได้ว่า

$$\begin{aligned} {}^nP_k &= n(n-1)\cdots(n-k+1) \\ &= n(n-1)\cdots(n-k+1)\frac{(n-k)!}{(n-k)!} \\ &= \frac{n!}{(n-k)!} \end{aligned}$$

<u>สรุป</u> การเรียงสับเปลี่ยนสิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง (ดังนั้น $k \le n$) สามารถเรียงสับเปลี่ยนได้ทั้งหมด

$${}^nP_k = \frac{n!}{(n-k)!} \quad \text{วิธี}$$

<u>หมายเหตุ</u>

1. $n!$ อ่านว่า n แฟคโทเรียล (n factorial) มีความหมายว่า

$$n! = n(n-1)(n-2)\ldots 1$$

2. $0! = 1$

3. การเรียงสับเปลี่ยนสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง จะสามารถทำได้ทั้งหมด

$${}^nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n! \text{ วิธี}$$

<u>**ตัวอย่างที่ 1.2.7**</u> ทีมฟุตบอลทีมหนึ่งรับสมัครนักกีฬา 2 ตำแหน่ง คือ กองกลาง และกองหลัง ปรากฏว่ามีนักกีฬาสมัครในตำแหน่งทั้ง 2 จำนวน 9 คน จงหาจำนวนวิธีในการจัดนักกีฬาเพื่อให้ประจำตำแหน่งทั้ง 2 ตำแหน่ง

วิธีทำ จำนวนวิธีในการจัดนักกีฬาเพื่อให้ประจำตำแหน่ง 2 ตำแหน่ง สำหรับนักกีฬาทั้งหมด 9 คน เปรียบเสมือนการเรียงสับเปลี่ยนสิ่งของ 2 สิ่ง จากที่มีอยู่ทั้งหมด 9 สิ่ง ซึ่งสามารถทำการเรียงสับเปลี่ยนได้ทั้งหมด

$${}^9P_2 = \frac{9!}{(9-2)!} = \frac{9!}{7!} = 72 \text{ วิธี}$$

ดังนั้น จำนวนวิธีในการจัดนักกีฬาเพื่อให้ประจำตำแหน่งทั้ง 2 ตำแหน่ง มีทั้งหมด 72 วิธี ■

<u>ข้อสังเกต</u>

สำหรับตัวอย่างที่ 1.2.7 ใช้การเรียงสับเปลี่ยนเพียงอย่างเดียวในการหาจำนวนสมาชิกของเหตุการณ์ ส่วนตัวอย่างที่ 1.2.8 จะเป็นตัวอย่างที่ต้องใช้การเรียงสับเปลี่ยน และหลักการคูณในการหาจำนวนสมาชิกของเหตุการณ์

<u>ตัวอย่างที่ 1.2.8</u> ต้องการสร้างคำภาษาอังกฤษที่ประกอบไปด้วยตัวอักษรที่แตกต่างกันทั้งหมด 7 ตัว โดยไม่คำนึงถึงความหมาย จากพยัญชนะ 4 ตัว และสระ 3 ตัว จะสร้างคำได้ทั้งหมดกี่วิธีโดยที่

(1) สระทั้ง 3 ตัว จะต้องอยู่แยกกันเสมอ

(2) สระทั้ง 3 ตัว จะต้องอยู่ติดกันเสมอ

วิธีทำ (1) การสร้างคำจากพยัญชนะ 4 ตัว และสระ 3 ตัว โดยที่สระทั้ง 3 ตัว จะต้องอยู่แยกกัน สามารถแบ่งออกได้เป็น 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1: จัดพยัญชนะ 4 ตัว เรียงเป็นแถวก่อน ซึ่งเปรียบได้กับการเรียงสับเปลี่ยนตำแหน่ง 4 ตำแหน่ง ดังนั้น สามารถทำการเรียงสับเปลี่ยนได้ทั้งหมด

$${}^{4}P_{4} = \frac{4!}{(4-4)!} = \frac{4!}{0!} = 24 \text{ วิธี}$$

ขั้นตอนที่ 2: จัดสระ 3 ตัว ใส่ในตำแหน่งระหว่างพยัญชนะทั้ง 4 ตัว ซึ่งมีตำแหน่งที่เป็นไปได้ทั้งหมด 5 ตำแหน่ง นั่นคือเปรียบได้กับการเรียงสับเปลี่ยนตำแหน่ง 3 ตำแหน่ง จากตำแหน่งทั้งหมด 5 ตำแหน่ง ดังนั้น สามารถทำการเรียงสับเปลี่ยนได้ทั้งหมด

$${}^{5}P_{3} = \frac{5!}{(5-3)!} = \frac{5!}{2!} = 60 \text{ วิธี}$$

จากหลักการคูณ สรุปได้ว่าจำนวนวิธีในการสร้างคำภาษาอังกฤษที่ประกอบไปด้วยตัวอักษรที่แตกต่างกันทั้งหมด 7 ตัว โดยไม่คำนึงถึงความหมาย จากพยัญชนะ 4 ตัว และสระ 3 ตัว โดยที่สระทั้ง 3 ตัว จะต้องอยู่แยกกันเสมอ สามารถทำได้ทั้งหมด $24 \times 60 = 1440$ วิธี

(2) การสร้างคำจากพยัญชนะ 4 ตัว และสระ 3 ตัว โดยที่สระทั้ง 3 ตัว จะต้องอยู่ติดกัน สามารถแบ่งออกได้เป็น 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1: พิจารณาเหมือนกับว่าสระทั้ง 3 ตัว รวมเป็นตัวเดียว นั่นคือ ทำการจัดตัวอักษร 5 ตัว (พยัญชนะ 4 ตัว สระคิดเหมือนมี 1 ตัว) เรียงเป็นแถว ซึ่งเปรียบได้กับการเรียงสับเปลี่ยนตำแหน่ง 5 ตำแหน่ง ดังนั้น สามารถทำการเรียงสับเปลี่ยนได้ทั้งหมด

$${}^{5}P_{5} = \frac{5!}{(5-5)!} = \frac{5!}{0!} = 120 \text{ วิธี}$$

ขั้นตอนที่ 2: สลับตำแหน่งสระ 3 ตัว ซึ่งจะทำให้เกิดคำใหม่ขึ้น เปรียบได้กับการเรียงสับเปลี่ยนตำแหน่ง 3 ตำแหน่ง ดังนั้น สามารถทำการเรียงสับเปลี่ยนได้ทั้งหมด

$${}^{3}P_{3} = \frac{3!}{(3-3)!} = \frac{3!}{0!} = 6 \text{ วิธี}$$

จากหลักการคูณ สรุปได้ว่าจำนวนวิธีในการสร้างคำภาษาอังกฤษที่ประกอบไปด้วยตัวอักษรที่แตกต่างกันทั้งหมด 7 ตัว โดยไม่คำนึงถึงความหมาย จากพยัญชนะ 4 ตัว และสระ 3 ตัว โดยที่สระทั้ง 3 ตัว จะต้องอยู่ติดกันเสมอ สามารถทำได้ทั้งหมด $120 \times 6 = 720$ วิธี ■

<u>หมายเหตุ</u>

สำหรับการเรียงสับเปลี่ยน นอกจากจะมีการเรียงสับเปลี่ยนเชิงเส้นสำหรับสิ่งของที่แตกต่างกันทั้งหมดแล้ว ยังมีการเรียงสับเปลี่ยนเชิงเส้นสำหรับสิ่งของที่ซ้ำกันบางส่วน และการเรียงสับเปลี่ยนเชิงวงกลมอีกด้วย ซึ่งในที่นี้จะกล่าวถึงเพียงแค่การเรียงสับเปลี่ยนเชิงเส้นซึ่งเป็นแนวความคิดพื้นฐานเท่านั้น ถ้าผู้อ่านต้องการทราบรายละเอียดเกี่ยวกับการเรียงสับเปลี่ยนเชิงเส้นสำหรับสิ่งของที่ซ้ำกันบางส่วน และการเรียงสับเปลี่ยนเชิงวงกลม สามารถอ่านเพิ่มเติมได้ในหนังสือเกี่ยวกับคอมบินาทอริก

1.2.4 การจัดหมู่

เหตุการณ์ที่เกี่ยวข้องกับการจัดสิ่งของเป็นกลุ่ม สามารถใช้การจัดหมู่มาช่วยในการนับจำนวนสมาชิกของเหตุการณ์ได้ ซึ่งจะทำให้การนับจำนวนสมาชิกของเหตุการณ์ลดความยุ่งยากซับซ้อนลง เมื่อเทียบกับการใช้หลักการคูณในการนับจำนวนสมาชิกของเหตุการณ์

<u>นิยามที่ 1.2.2</u> ***การจัดหมู่ (combination)*** หมายถึง การนำสิ่งของบางส่วนหรือทั้งหมดมาจัดเป็นกลุ่ม โดย "ไม่คำนึงถึงลำดับ" นั่นคือ ถึงแม้จะมีการสลับลำดับหรือตำแหน่ง ก็ยังถือว่าเป็นการจัดหมู่ที่ได้สมาชิกในกลุ่มแบบเดิม ไม่ทำให้เกิดการจัดหมู่แบบใหม่ที่แตกต่างจากเดิม

<u>หมายเหตุ</u>

การจัดหมู่ บางครั้งอาจเรียกว่า การจัดกลุ่ม การเปลี่ยนหมู่ หรือการเปลี่ยนกลุ่ม

<u>ตัวอย่างที่ 1.2.9</u> ถ้ามีตัวอักษรภาษาอังกฤษทั้งหมด 3 ตัว คือ A, B และ C จะสามารถทำการจัดหมู่ได้ทั้งหมดกี่วิธี อะไรบ้างเมื่อ

(1) ต้องการจัดหมู่ตัวอักษร 1 ตัว

(2) ต้องการจัดหมู่ตัวอักษร 2 ตัว

(3) ต้องการจัดหมู่ตัวอักษร 3 ตัว

วิธีทำ (1) การจัดหมู่ตัวอักษร 1 ตัว จากตัวอักษรทั้งหมด 3 ตัว (A, B และ C) ทำได้ทั้งหมด 3 วิธี ดังนี้ A, B และ C

(2) การจัดหมู่ตัวอักษร 2 ตัว จากตัวอักษรทั้งหมด 3 ตัว (A, B และ C) ทำได้ทั้งหมด 3 วิธี ดังนี้ AB, AC และ BC

(3) การจัดหมู่ตัวอักษร 3 ตัว จากตัวอักษรทั้งหมด 3 ตัว (A, B และ C) ทำได้ทั้งหมด 1 วิธี ดังนี้ ABC ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 1.2.9 จะสังเกตได้ว่าการจัดหมู่ตัวอักษร 3 ตัว จากตัวอักษรทั้งหมด 3 ตัว ABC และ ACB ถึงแม้ว่าลำดับของตัวอักษรจะแตกต่างกัน แต่มีตัวอักษรเหมือนกันทั้งหมด ทำให้การจัดหมู่ทั้ง 2 วิธีนี้เหมือนกัน นั่นคือ สำหรับการจัดหมู่ลำดับไม่มีความสำคัญ ถึงแม้จะมีการสลับลำดับของสิ่งของที่อยู่ในกลุ่ม แต่ถ้าสมาชิกในกลุ่มยังคงเหมือนเดิม จะถือว่าเป็นการจัดหมู่แบบเดิม

พิจารณาการจัดหมู่สิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง (ดังนั้น $k \le n$)

กำหนดให้ ${}^{n}C_k$ หรือ $\binom{n}{k}$ แทนจำนวนวิธีการจัดหมู่สิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง ถ้าพิจารณาการเรียงสับเปลี่ยนสิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง สามารถแบ่งออกได้เป็น 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1: จัดหมู่สิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง สามารถจัดหมู่สิ่งของได้ทั้งหมด ${}^{n}C_k$ หรือ $\binom{n}{k}$ วิธี

ขั้นตอนที่ 2: นำสิ่งของในขั้นตอนที่ 1 ซึ่งมีทั้งหมด k สิ่ง มาทำการเรียงสับเปลี่ยนทั้งหมด สามารถทำได้ทั้งหมด ${}^{k}P_k = k!$ วิธี

จากหลักการคูณ สามารถสรุปได้ว่าจำนวนวิธีการเรียงสับเปลี่ยนสิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง สามารถทำได้ทั้งหมด ${}^{n}C_k \times k!$ หรือ $\binom{n}{k} \times k!$ วิธี นั่นคือ

$${}^{n}P_k = {}^{n}C_k \times k! = \binom{n}{k} \times k!$$

หรือ

$${}^{n}C_k = \binom{n}{k} = \frac{{}^{n}P_k}{k!} = \frac{n!}{k!(n-k)!}$$

<u>สรุป</u> การจัดหมู่สิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง (ดังนั้น $k \le n$) สามารถจัดหมู่ได้ทั้งหมด

$${}^{n}C_k = \binom{n}{k} = \frac{n!}{k!(n-k)!} \quad \text{วิธี}$$

<u>หมายเหตุ</u>

ความสัมพันธ์ระหว่างการเรียงสับเปลี่ยนและการจัดหมู่สิ่งของ k สิ่ง จากสิ่งของที่แตกต่างกันทั้งหมด n สิ่ง (ดังนั้น $k \le n$) เป็นไปตามสมการต่อไปนี้

$${}^{n}C_k = \binom{n}{k} = \frac{{}^{n}P_k}{k!}$$

<u>ตัวอย่างที่ 1.2.10</u> ถ้ามีหุ้นของบริษัทอยู่ 7 บริษัท คือ A, B, C, D, E, F และ G ต้องการเลือกลงทุนในหุ้นของบริษัท 4 บริษัท จะสามารถเลือกลงทุนได้ทั้งหมดกี่วิธี

วิธีทำ การเลือกลงทุนในหุ้นของบริษัท 4 บริษัท จากทั้งหมด 7 บริษัท เปรียบเสมือนการจัดหมู่สิ่งของ 4 สิ่ง จากสิ่งของทั้งหมด 7 สิ่ง ซึ่งสามารถจัดหมู่ได้ทั้งหมด

$$ {}^{7}C_{4} = \frac{7!}{4!(7-4)!} = \frac{7!}{4!3!} = 35 \text{ วิธี} $$

ดังนั้น การเลือกลงทุนในหุ้นของบริษัท 4 บริษัท จาก 7 บริษัท สามารถเลือกลงทุนได้ทั้งหมด 35 วิธี ■

<u>ข้อสังเกต</u>

การหาจำนวนสมาชิกของเหตุการณ์สำหรับตัวอย่างที่ 1.2.10 ใช้การจัดหมู่เพียงอย่างเดียว ส่วนตัวอย่างที่ 1.2.11 และ 1.2.12 จะเป็นตัวอย่างที่ต้องใช้การจัดหมู่ หลักการบวก และหลักการคูณในการหาจำนวนสมาชิกของเหตุการณ์

<u>ตัวอย่างที่ 1.2.11</u> ร้านขายเสื้อผ้าแห่งหนึ่งมีเสื้อที่แตกต่างกันทั้งหมด 12 ตัว ประกอบไปด้วยเสื้อแขนสั้น 5 ตัว และเสื้อแขนยาว 7 ตัว ถ้าลูกค้าต้องการเลือกซื้อเสื้ออย่างสุ่มจำนวน 6 ตัว จงหาจำนวนวิธีที่ลูกค้าจะเลือกได้เสื้อแขนสั้นอย่างน้อย 4 ตัว

วิธีทำ เหตุการณ์ที่ลูกค้าจะเลือกได้เสื้อแขนสั้นอย่างน้อย 4 ตัว แบ่งออกเป็น 2 กรณี ที่ไม่เกิดพร้อมกันดังนี้

กรณีที่ 1: เหตุการณ์ที่ลูกค้าเลือกได้เสื้อแขนสั้น 4 ตัว สามารถแบ่งได้เป็น 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1: ลูกค้าเลือกเสื้อแขนสั้น 4 ตัว จากเสื้อแขนสั้นทั้งหมด 5 ตัว เปรียบเสมือนการจัดหมู่สิ่งของ 4 สิ่ง จากสิ่งของทั้งหมด 5 สิ่ง ซึ่งสามารถจัดหมู่ได้ทั้งหมด

$$ {}^{5}C_{4} = \frac{5!}{4!(5-4)!} = \frac{5!}{4!1!} = 5 \text{ วิธี} $$

ขั้นตอนที่ 2: ลูกค้าเลือกเสื้อแขนยาว 2 ตัว จากเสื้อแขนยาวทั้งหมด 7 ตัว เปรียบเสมือนการจัดหมู่สิ่งของ 2 สิ่ง จากสิ่งของทั้งหมด 7 สิ่ง ซึ่งสามารถจัดหมู่ได้ทั้งหมด

$$ {}^{7}C_{2} = \frac{7!}{2!(7-2)!} = \frac{7!}{2!5!} = 21 \text{ วิธี} $$

จากหลักการคูณ สรุปได้ว่าสำหรับการเลือกซื้อเสื้ออย่างสุ่มจำนวน 6 ตัว จำนวนวิธีที่ลูกค้าจะเลือกได้เสื้อแขนสั้น 4 ตัว เป็นไปได้ทั้งหมด $5 \times 21 = 105$ วิธี

กรณีที่ 2: เหตุการณ์ที่ลูกค้าเลือกได้เสื้อแขนสั้น 5 ตัว สามารถแบ่งออกได้เป็น 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1: ลูกค้าเลือกเสื้อแขนสั้น 5 ตัว จากเสื้อแขนสั้นทั้งหมด 5 ตัว เปรียบเสมือนการจัดหมู่สิ่งของ 5 สิ่ง จากสิ่งของทั้งหมด 5 สิ่ง ซึ่งสามารถจัดหมู่ได้ทั้งหมด

$$ {}^{5}C_{5} = \frac{5!}{5!(5-5)!} = \frac{5!}{5!0!} = 1 \text{ วิธี} $$

ขั้นตอนที่ 2: ลูกค้าเลือกเสื้อแขนยาว 1 ตัว จากเสื้อแขนยาวทั้งหมด 7 ตัว เปรียบเสมือนการจัดหมู่สิ่งของ 1 สิ่ง จากสิ่งของทั้งหมด 7 สิ่ง ซึ่งสามารถจัดหมู่ได้ทั้งหมด

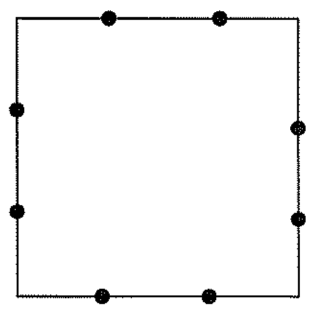
$${}^{7}C_1 = \frac{7!}{1!(7-1)!} = \frac{7!}{1!6!} = 7 \text{ วิธี}$$

จากหลักการคูณ สรุปได้ว่าสำหรับการเลือกซื้อเสื้ออย่างสุ่มจำนวน 6 ตัว จำนวนวิธีที่ลูกค้าจะเลือกได้เสื้อแขนสั้น 5 ตัว เป็นไปได้ทั้งหมด $1 \times 7 = 7$ วิธี

จากหลักการบวก สรุปได้ว่าสำหรับการเลือกซื้อเสื้ออย่างสุ่มจำนวน 6 ตัว จำนวนวิธีที่ลูกค้าจะเลือกได้เสื้อแขนสั้นอย่างน้อย 4 ตัว เป็นไปได้ทั้งหมด $105 + 7 = 112$ วิธี ■

<u>**ตัวอย่างที่ 1.2.12**</u> รูปสี่เหลี่ยมรูปหนึ่งมีจุดอยู่บนด้านของรูปสี่เหลี่ยมด้านละ 2 จุด ดังรูปที่ 1.2.6 ถ้าต้องการสร้างรูปเหลี่ยมบรรจุภายในรูปสี่เหลี่ยมดังกล่าว โดยมีจุดมุมเป็นจุดเหล่านี้ จงหาจำนวนวิธีในการสร้างรูปเหลี่ยมทั้งหมดที่เป็นไปได้

วิธีทำ รูปสี่เหลี่ยมที่มีจุด 2 จุด อยู่บนด้านแต่ละด้าน แสดงได้ในรูปที่ 1.2.6



รูปที่ 1.2.6 รูปสี่เหลี่ยมที่มีจุด 2 จุด อยู่บนด้านแต่ละด้าน

จำนวนรูปเหลี่ยมทั้งหมดที่สร้างได้จากจุด 8 จุด สามารถแบ่งออกได้เป็น 6 กรณี ที่ไม่เกิดร่วมกันดังนี้

กรณีที่ 1: การสร้างรูปสามเหลี่ยม เนื่องจากลำดับในการเลือกจุดไม่สำคัญ จึงถือว่าเป็นการจัดหมู่จุด 3 จุด จากทั้งหมด 8 จุด ดังนั้น สร้างรูปสามเหลี่ยมได้ทั้งหมด

$${}^{8}C_3 = \frac{8!}{3!(8-3)!} = \frac{8!}{3!5!} = 56 \text{ วิธี}$$

กรณีที่ 2: การสร้างรูปสี่เหลี่ยม ถือว่าเป็นการจัดหมู่จุด 4 จุด จากทั้งหมด 8 จุด ดังนั้น สร้างรูปสี่เหลี่ยมได้ทั้งหมด

$${}^{8}C_4 = \frac{8!}{4!(8-4)!} = \frac{8!}{4!4!} = 70 \text{ วิธี}$$

กรณีที่ 3: การสร้างรูปห้าเหลี่ยม ถือว่าเป็นการจัดหมู่จุด 5 จุด จากทั้งหมด 8 จุด ดังนั้น สร้างรูปห้าเหลี่ยมได้ทั้งหมด

$${}^{8}C_5 = \frac{8!}{5!(8-5)!} = \frac{8!}{5!3!} = 56 \text{ วิธี}$$

กรณีที่ 4: การสร้างรูปหกเหลี่ยม ถือว่าเป็นการจัดหมู่จุด 6 จุด จากทั้งหมด 8 จุด ดังนั้น สร้างรูปหกเหลี่ยมได้ทั้งหมด

$$^{8}C_{6}=\frac{8!}{6!(8-6)!}=\frac{8!}{6!2!}=28 \text{ วิธี}$$

กรณีที่ 5: การสร้างรูปเจ็ดเหลี่ยม ถือว่าเป็นการจัดหมู่จุด 7 จุด จากทั้งหมด 8 จุด ดังนั้น สร้างรูปเจ็ดเหลี่ยมได้ทั้งหมด

$$^{8}C_{7}=\frac{8!}{7!(8-7)!}=\frac{8!}{7!1!}=8 \text{ วิธี}$$

กรณีที่ 6: การสร้างรูปแปดเหลี่ยม ถือว่าเป็นการจัดหมู่จุด 8 จุด ดังนั้น สร้างรูปแปดเหลี่ยมได้ทั้งหมด

$$^{8}C_{8}=\frac{8!}{8!(8-8)!}=\frac{8!}{8!0!}=1 \text{ วิธี}$$

จากหลักการบวก สรุปได้ว่าสามารถสร้างรูปเหลี่ยมจากจุด 8 จุด ที่อยู่บนเส้นรอบรูปสี่เหลี่ยมได้ทั้งหมด $56+70+56+28+8+1=219$ วิธี ■

1.3 เซต

เนื่องจากปริภูมิตัวอย่างและเหตุการณ์สามารถนิยามได้ด้วยเซต ดังนั้น เซตจึงเป็นความรู้พื้นฐานที่สำคัญในการศึกษาเกี่ยวกับปริภูมิตัวอย่างและเหตุการณ์ ในหัวข้อนี้จึงได้ทบทวนนิยาม และสมบัติที่สำคัญเกี่ยวกับเซตที่จำเป็นต้องใช้ต่อไป

<u>นิยามที่ 1.3.1</u> กำหนดให้ A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง นั่นคือ $A,B\subset S$ ดังนั้น

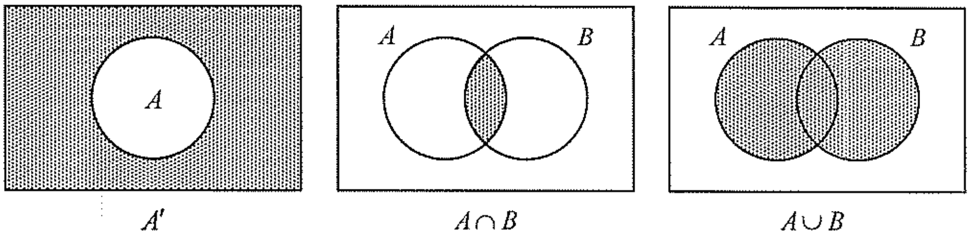
คอมพลีเมนต์ (complement) ของเหตุการณ์ A หมายถึง เซตของสมาชิกทั้งหมดของปริภูมิตัวอย่าง S ที่ไม่เป็นสมาชิกของเหตุการณ์ A เขียนแทนได้ด้วยสัญลักษณ์ A'

อินเตอร์เซกชัน (intersection) ของเหตุการณ์ A และ B หมายถึง เซตของสมาชิกทั้งหมดที่เป็นสมาชิกของเหตุการณ์ A และ B เขียนแทนได้ด้วยสัญลักษณ์ $A\cap B$

ยูเนียน (union) ของเหตุการณ์ A และ B หมายถึง เซตของสมาชิกทั้งหมดที่เป็นสมาชิกของเหตุการณ์ A หรือ B เขียนแทนได้ด้วยสัญลักษณ์ $A\cup B$

เหตุการณ์ที่ไม่เกิดร่วมกัน (mutually disjoint event) เหตุการณ์ $A_1,\ldots,A_n$ จะเรียกว่าเป็นเหตุการณ์ที่ไม่เกิดร่วมกัน เมื่อ $A_i\cap A_j=\varnothing,\forall i\neq j$ โดยที่ $i,j=1,2,\ldots,n$

แผนภาพเวนน์และออยเลอร์แสดงเหตุการณ์ที่เกิดจากการกระทำของเซต แสดงได้ในรูปที่ 1.3.1 โดยที่ปริภูมิตัวอย่างแทนได้ด้วยรูปสี่เหลี่ยมผืนผ้า ส่วนเหตุการณ์แทนได้ด้วยรูปวงกลม และบริเวณที่แรเงาหมายถึงเหตุการณ์ A', $A\cap B$ และ $A\cup B$ ตามลำดับ



รูปที่ 1.3.1 แผนภาพเวนน์และออยเลอร์ของเหตุการณ์ A' , $A \cap B$ และ $A \cup B$ ตามลำดับ

สมบัติที่สำคัญเกี่ยวกับเซต

กำหนดให้ A, B และ C เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง นั่นคือ $A, B, C \subset S$ (เหตุการณ์ A, B และ C เปรียบได้กับเซต A, B และ C ส่วนปริภูมิตัวอย่างเปรียบได้กับเอกภพสัมพัทธ์หรือ S) ดังนั้น

1. $A \cap B = B \cap A$
2. $A \cup B = B \cup A$
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
5. $(A \cap B)' = A' \cup B'$
6. $(A \cup B)' = A' \cap B'$
7. $A \cap A' = \varnothing$
8. $A \cup A' = S$
9. $A - B = A \cap B'$

<u>หมายเหตุ</u>

สำหรับหัวข้อที่ 1.3 ได้ทบทวนความรู้บางส่วนเรื่องเซต ถ้าผู้อ่านต้องการทราบรายละเอียดเพิ่มเติมเรื่องเซตสามารถอ่านได้ในหนังสือเกี่ยวกับเซต

1.4 ความน่าจะเป็นและสมบัติของความน่าจะเป็น

ในหัวข้อนี้จะกล่าวถึงความหมายของความน่าจะเป็น การคำนวณหาค่าความน่าจะเป็นของเหตุการณ์ต่าง ๆ ในปริภูมิตัวอย่าง สัจพจน์ของความน่าจะเป็น และสมบัติที่สำคัญของความน่าจะเป็น

1.4.1 ความน่าจะเป็น

<u>**นิยามที่ 1.4.1**</u> ***ความน่าจะเป็น (probability)*** ของเหตุการณ์ หมายถึง ค่าที่บอกว่าเหตุการณ์นั้นจะมีโอกาสเกิดขึ้นมากน้อยเพียงใด ความน่าจะเป็นของเหตุการณ์ A เขียนแทนได้ด้วยสัญลักษณ์ $P(A)$

<u>ตัวอย่างที่ 1.4.1</u> ตัวอย่างความน่าจะเป็นของเหตุการณ์

(1) ความน่าจะเป็นที่ราคาหุ้นของบริษัท A จะเพิ่มขึ้นในสัปดาห์หน้ามีค่าเท่ากับ 0.6

(2) ความน่าจะเป็นที่ฝนจะตกในวันพรุ่งนี้มีค่าเท่ากับ 0.3

<u>ข้อสังเกต</u>

1. ความน่าจะเป็นของเหตุการณ์ใด ๆ มีค่าตั้งแต่ 0 ถึง 1
2. ความน่าจะเป็นของเหตุการณ์มีค่าเท่ากับ 1 หมายความว่า เหตุการณ์นั้นเกิดขึ้นแน่นอน
3. ความน่าจะเป็นของเหตุการณ์มีค่าเท่ากับ 0 หมายความว่า เหตุการณ์นั้นไม่เกิดขึ้นแน่นอน
4. ความน่าจะเป็นของเหตุการณ์มีค่าเท่ากับ 0.5 หมายความว่า เหตุการณ์นั้นมีโอกาสเกิดขึ้น 50%
5. ความน่าจะเป็นบางครั้งแทนได้ด้วยร้อยละหรือเปอร์เซ็นต์

1.4.2 การคำนวณหาค่าความน่าจะเป็น

การคำนวณหาค่าความน่าจะเป็น แบ่งออกได้เป็น 2 วิธี ดังนี้

1. วิธีคลาสสิก (classical method) การหาค่าความน่าจะเป็นด้วยวิธีคลาสสิก สามารถหาค่าความน่าจะเป็นของเหตุการณ์ A ได้ดังนี้

$$P(A)=\frac{n(A)}{n(S)}$$

เมื่อ $n(A)$ และ $n(S)$ หมายถึง จำนวนสมาชิกของเหตุการณ์ A และปริภูมิตัวอย่าง S ตามลำดับ

สำหรับการหาค่าความน่าจะเป็นด้วยวิธีคลาสสิกมีข้อกำหนดว่าสมาชิกแต่ละตัวในปริภูมิตัวอย่างจะต้องมีโอกาสเกิดขึ้นเท่า ๆ กัน ดังนั้น หลักการบวก หลักการคูณ การเรียงสับเปลี่ยน และการจัดหมู่ ที่ได้กล่าวถึงในหัวข้อที่ 1.2 จึงสามารถนำมาใช้ในการนับจำนวนสมาชิกของเหตุการณ์หรือปริภูมิตัวอย่างได้

2. วิธีใช้ความถี่สัมพัทธ์ (relative frequency approach) วิธีนี้อาศัยความถี่ของข้อมูลซึ่งเกิดขึ้นจริงในอดีตจำนวนมาก เพื่อนำมาหาค่าความน่าจะเป็นในอนาคต โดยที่ความน่าจะเป็นของเหตุการณ์ A หาค่าได้ดังนี้

$$P(A)=\frac{f}{n}$$

เมื่อ f หมายถึง จำนวนครั้งที่เกิดเหตุการณ์ A และ n หมายถึง จำนวนครั้งของการเกิดเหตุการณ์ทั้งหมด

<u>ตัวอย่างที่ 1.4.2</u> ประธานค่ายอาสาฯ ค่ายหนึ่ง ต้องการรับสมัครคนเพื่อจัดค่ายจำนวน 7 คน แต่มีผู้สมัครจำนวน 11 คน เป็นชาย 6 คน และเป็นหญิง 5 คน ถ้าเลือกคนเพื่อจัดค่ายโดยการสุ่ม จงหาความน่าจะเป็นที่เลือกคนเพื่อมาจัดค่ายเป็นชาย 4 คน และหญิง 3 คน

วิธีทำ ให้ A แทนเหตุการณ์ที่เลือกคนเพื่อจัดค่ายเป็นชาย 4 คน และหญิง 3 คน เนื่องจากสมาชิกในปริภูมิตัวอย่างของการเลือกคนเพื่อจัดค่าย 7 คน จากผู้สมัครทั้งหมด 11 คน มีโอกาสเกิดขึ้นเท่า ๆ กัน ดังนั้น จึงสามารถใช้วิธีคลาสสิกในการหาค่าความน่าจะเป็นของเหตุการณ์ A ได้ดังนี้

$$P(A)=\frac{n(A)}{n(S)}$$

จะสังเกตได้ว่า การหาจำนวนสมาชิกของเหตุการณ์ A และปริภูมิตัวอย่าง S จำเป็นต้องใช้หลักการนับพื้นฐานเข้ามาช่วย ซึ่งทำได้ดังนี้

$$P(A)=\frac{{}^{6}C_{4}\,{}^{5}C_{3}}{{}^{11}C_{7}}=\frac{\dfrac{6!}{4!(6-4)!}\cdot\dfrac{5!}{3!(5-3)!}}{\dfrac{11!}{7!(11-7)!}}=\frac{150}{330}=0.4545$$

ดังนั้น ความน่าจะเป็นที่เลือกคนเพื่อจัดค่ายเป็นชาย 4 คน และหญิง 3 คน มีค่าเท่ากับ 0.4545 ■

<u>ตัวอย่างที่ 1.4.3</u> การศึกษาความสัมพันธ์ระหว่างประเภทของหอพัก ชั้นปี และเพศของนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ของมหาวิทยาลัยแห่งหนึ่งซึ่งมีจำนวน 160 คน ได้ข้อมูลจำแนกตามเพศ ชั้นปี และประเภทของหอพัก ดังนี้

ชั้นปีที่	หอพักในมหาวิทยาลัย		หอพักนอกมหาวิทยาลัย	
	ชาย	หญิง	ชาย	หญิง
1	12	12	1	1
2	10	7	20	17
3	9	8	15	12
4	5	2	11	10
อื่น ๆ	2	1	5	0
รวม	38	30	52	40

(1) ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน จงหาความน่าจะเป็นที่นักศึกษาที่สุ่มมาจะเป็นนักศึกษาชั้นปีที่ 1 ที่อยู่หอพักในมหาวิทยาลัย

(2) ถ้าสุ่มนักศึกษากลุ่มนี้ที่อยู่หอพักนอกมหาวิทยาลัยมา 1 คน จงหาความน่าจะเป็นที่นักศึกษาที่สุ่มมาจะเป็นนักศึกษาชาย

(3) ถ้าสุ่มนักศึกษากลุ่มนี้ที่เป็นนักศึกษาชั้นปีที่ 4 มา 1 คน จงหาความน่าจะเป็นที่นักศึกษาที่สุ่มมาจะอยู่หอพักนอกมหาวิทยาลัย

วิธีทำ (1) ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน กำหนดให้ A แทนเหตุการณ์ที่สุ่มได้นักศึกษาชั้นปีที่ 1 ที่อยู่หอพักในมหาวิทยาลัย เนื่องจากสุ่มนักศึกษากลุ่มนี้มา 1 คน ดังนั้น จำนวนสมาชิกของปริภูมิตัวอย่าง จึงมีค่าเท่ากับจำนวนของนักศึกษากลุ่มนี้ทั้งหมด ซึ่งจะได้ว่า

$$P(A)=\frac{24}{160}=0.15$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน ความน่าจะเป็นที่นักศึกษาที่สุ่มมาจะเป็นนักศึกษาชั้นปีที่ 1 ที่อยู่หอพักในมหาวิทยาลัยมีค่าเท่ากับ 0.15

(2) ถ้าสุ่มนักศึกษากลุ่มนี้ที่อยู่หอพักนอกมหาวิทยาลัยมา 1 คน กำหนดให้ B แทนเหตุการณ์ที่สุ่มได้นักศึกษาชาย เนื่องจากสุ่มนักศึกษากลุ่มนี้ที่อยู่หอพักนอกมหาวิทยาลัยมา 1 คน ดังนั้น จำนวนสมาชิกของปริภูมิตัวอย่าง จึงมีค่าเท่ากับจำนวนนักศึกษากลุ่มนี้ทั้งหมดที่อยู่หอพักนอกมหาวิทยาลัย ซึ่งจะได้ว่า

$$P(B) = \frac{52}{92} = 0.5652$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้ที่อยู่หอพักนอกมหาวิทยาลัยมา 1 คน ความน่าจะเป็นที่นักศึกษาที่สุ่มมาจะเป็นนักศึกษาชายมีค่าเท่ากับ 0.5652

(3) ถ้าสุ่มนักศึกษากลุ่มนี้ที่เป็นนักศึกษาชั้นปีที่ 4 มา 1 คน กำหนดให้ C แทนเหตุการณ์ที่สุ่มได้นักศึกษาที่อยู่หอพักนอกมหาวิทยาลัย เนื่องจากสุ่มนักศึกษากลุ่มนี้ที่เป็นนักศึกษาชั้นปีที่ 4 มา 1 คน ดังนั้น จำนวนสมาชิกของปริภูมิตัวอย่าง จึงมีค่าเท่ากับจำนวนนักศึกษากลุ่มนี้ทั้งหมดที่เป็นนักศึกษาชั้นปีที่ 4 ซึ่งจะได้ว่า

$$P(C) = \frac{21}{28} = 0.75$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้ที่เป็นนักศึกษาชั้นปีที่ 4 มา 1 คน ความน่าจะเป็นที่นักศึกษาที่สุ่มมาจะอยู่หอพักนอกมหาวิทยาลัยมีค่าเท่ากับ 0.75 ■

1.4.3 สัจพจน์ของความน่าจะเป็น

นิยามที่ 1.4.2 กำหนดให้ S เป็นปริภูมิตัวอย่าง ความน่าจะเป็น P เป็นฟังก์ชันที่ส่งจากปริภูมิตัวอย่างไปยังเซตของจำนวนจริง ซึ่งสอดคล้องกับเงื่อนไขต่อไปนี้

(1) ถ้า A เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A \subset S$) แล้ว

$$P(A) \geq 0$$

(2) สำหรับปริภูมิตัวอย่าง S แล้ว

$$P(S) = 1$$

(3) ถ้า A_i โดยที่ $i = 1, 2, \ldots$ เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A_i \subset S$ สำหรับทุก $i = 1, 2, \ldots$) ซึ่งเป็นเหตุการณ์ที่ไม่เกิดร่วมกัน (นั่นคือ $A_i \cap A_j = \varnothing$, สำหรับทุก $i \neq j$ โดยที่ $i, j = 1, \ldots$) แล้ว

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

ข้อสังเกต จากสัจพจน์ของความน่าจะเป็นข้อที่ (3) จะได้ว่า

1. ถ้า A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B \subset S$) และ A, B เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน (นั่นคือ $A \cap B = \varnothing$) แล้ว

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

2. ถ้า A_1, $A_2,\ldots,$ A_n เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A_1,A_2,\ldots A_n \subset S$) ซึ่งเป็นเหตุการณ์ที่ไม่เกิดร่วมกัน (นั่นคือ $A_i \cap A_j = \varnothing$ สำหรับทุก $i \neq j$ โดยที่ $i,j=1,\ldots,n$) แล้ว

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1)+P(A_2)+\ldots+P(A_n)$$

1.4.4 สมบัติที่สำคัญของความน่าจะเป็น

ในหัวข้อนี้จะกล่าวถึงสมบัติที่สำคัญของความน่าจะเป็นที่ช่วยให้การคำนวณหาค่าความน่าจะเป็นสะดวกและง่ายขึ้น สมบัติต่าง ๆ เหล่านี้เป็นผลมาจากสัจพจน์ของความน่าจะเป็นในหัวข้อที่ 1.4.3 ซึ่งสมบัติที่สำคัญของความน่าจะเป็นประกอบไปด้วย

<u>ข้อที่ 1</u>: ถ้า A เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A \subset S$) แล้ว

$$P(A') = 1 - P(A)$$

<u>พิสูจน์</u> เนื่องจาก $S = A \cup A'$ พิจารณา

$$P(S) = P(A \cup A')$$

และ $A \cap A' = \varnothing$ ดังนั้น เหตุการณ์ A และ A' เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน จากสัจพจน์ของความน่าจะเป็นข้อที่ (3) จึงได้ว่า $P(A \cup A') = P(A) + P(A')$ นั่นคือ

$$P(S) = P(A) + P(A')$$

จากสัจพจน์ของความน่าจะเป็นข้อที่ (2) ทำให้ได้ว่า

$$1 = P(A) + P(A')$$

ดังนั้น

$$P(A') = 1 - P(A)$$

<u>ข้อที่ 2</u>: ถ้า $\varnothing$ เป็นเหตุการณ์ที่เป็นไปไม่ได้ แล้ว

$$P(\varnothing) = 0$$

<u>พิสูจน์</u> เนื่องจาก $\varnothing = S'$ พิจารณา

$$P(\varnothing) = P(S')$$

จากสมบัติของความน่าจะเป็นข้อที่ 1 จะได้ว่า $P(S') = 1 - P(S)$ ทำให้ได้ว่า

$$P(\varnothing) = 1 - P(S)$$

และจากสัจพจน์ของความน่าจะเป็นข้อที่ (2) ดังนั้น

$$P(\varnothing) = 1 - P(S) = 1 - 1 = 0$$

<u>ข้อที่ 3</u>: ถ้า A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B \subset S$) แล้ว

$$P(A) = P(A \cap B) + P(A \cap B')$$

<u>พิสูจน์</u> เนื่องจาก $A = A\cap S = A\cap(B\cup B') = (A\cap B)\cup(A\cap B')$ พิจารณา

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

และ $(A\cap B)\cap(A\cap B') = \varnothing$ ดังนั้น เหตุการณ์ $A\cap B$ และ $A\cap B'$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน จากสัจพจน์ของความน่าจะเป็นข้อที่ (3) จึงได้ว่า $P((A\cap B)\cup(A\cap B')) = P(A\cap B) + P(A\cap B')$ ดังนั้น

$$P(A) = P(A\cap B) + P(A\cap B')$$

<u>ข้อที่ 4</u>: ถ้า A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B\subset S$) แล้ว

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

<u>พิสูจน์</u> จากสมบัติของความน่าจะเป็นข้อที่ 3 และ $A\cap B' = A-B$ ดังนั้น

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

นั่นคือ

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

<u>ข้อที่ 5</u>: ถ้า A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B\subset S$) แล้ว

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

<u>พิสูจน์</u> เนื่องจาก $A\cup B = (A\cup B)\cap S = (A\cup B)\cap(A\cup A') = A\cup(B\cap A')$ พิจารณา

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

และ $A\cap(B\cap A') = \varnothing$ ดังนั้น เหตุการณ์ A และ $B\cap A'$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน จากสัจพจน์ของความน่าจะเป็นข้อที่ (3) จึงได้ว่า $P(A\cup(B\cap A')) = P(A) + P(B\cap A')$ นั่นคือ

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

จากสมบัติของความน่าจะเป็นข้อที่ (3) $P(B) = P(B\cap A) + P(B\cap A')$ หรือ $P(B\cap A') = P(B) - P(B\cap A)$ ดังนั้น

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

<u>ข้อที่ 6</u>: ถ้า A, B และ C เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B, C\subset S$) แล้ว

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

<u>พิสูจน์</u> พิจารณา $A\cup B\cup C = (A\cup B)\cup C$ โดยใช้สมบัติของความน่าจะเป็นข้อที่ 5 จะได้ว่า

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

เนื่องจากสมบัติของความน่าจะเป็นข้อที่ 5 $P(A\cup B) = P(A) + P(B) - P(A\cap B)$ ทำให้ได้ว่า

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

จาก $(A\cup B)\cap C = (A\cap C)\cup(B\cap C)$ และสมบัติของความน่าจะเป็นข้อที่ 5 จะได้ว่า

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

ดังนั้น สามารถสรุปได้ว่า

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

<u>ข้อที่ 7</u>: ถ้า $A_1, A_2, \ldots A_n$ เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่างที่ไม่เกิดร่วมกัน (นั่นคือ $A_1, A_2, \ldots A_n \subset S$ และ $A_i\cap A_j=\varnothing, \forall i\neq j$ โดยที่ $i,j=1,\ldots,n$) และ $\bigcup_{i=1}^{n}A_i=S$ แล้ว

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

<u>พิสูจน์</u> จากสัจพจน์ของความน่าจะเป็นข้อที่ (3) จะได้ว่า

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

เนื่องจาก $\bigcup_{i=1}^{n}A_i=S$ และสัจพจน์ของความน่าจะเป็นข้อที่ (2) จึงได้ว่า $P\left(\bigcup_{i=1}^{n}A_i\right)=P(S)=1$ ดังนั้น

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

<u>ข้อที่ 8</u>: ถ้า A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง โดยที่ $B\subset A\subset S$ แล้ว

$$P(B)\le P(A)$$

<u>พิสูจน์</u> พิจารณา $A=A\cap S=A\cap(B\cup B')=(A\cap B)\cup(A\cap B')$ และเนื่องจาก $B\subset A$ ดังนั้น $A\cap B=B$ นั่นคือ $A=B\cup(A\cap B')$ จากนั้นพิจารณา

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

และ $B\cap(A\cap B')=\varnothing$ ดังนั้น เหตุการณ์ B และ $A\cap B'$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน จากสัจพจน์ของความน่าจะเป็นข้อที่ (3) จึงได้ว่า $P(B\cup(A\cap B'))=P(B)+P(A\cap B')$ นั่นคือ

$$P(A)=P(B)+P(A\cap B')$$

เนื่องจากสัจพจน์ของความน่าจะเป็นข้อที่ (1) จึงได้ว่า $P(A\cap B')\ge 0$ ดังนั้น สามารถสรุปได้ว่า

$$P(B)\le P(A)$$

<u>ข้อที่ 9</u>: ถ้า A เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A\subset S$) แล้ว

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

<u>พิสูจน์</u> จากสมบัติของความน่าจะเป็นข้อที่ 8 และ $A \subset S$ จะได้ว่า

$$P(A) \le P(S)$$

เนื่องจากสัจพจน์ของความน่าจะเป็นข้อที่ (2) $P(S) = 1$ จะได้ว่า

$$P(A) \le 1$$

และจากสัจพจน์ของความน่าจะเป็นข้อที่ (1) ดังนั้น

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

<u>ตัวอย่างที่ 1.4.4</u> กำหนดให้ A, B และ C เป็นเหตุการณ์ในปริภูมิตัวอย่าง โดยที่ $P(A) = 0.3$, $P(B) = 0.2$ และ $P(A \cap B) = 0.1$ จงหาความน่าจะเป็นต่อไปนี้

(1) $P(A')$

(2) $P(A \cup B)$

(3) $P[(A \cup B)']$

(4) $P[(A \cup B)' \cup B]$

(5) $P(A - B)$

(6) $P(A' \cap B)$

(7) $P(A' \cup B)$

วิธีทำ (1) จากสมบัติของความน่าจะเป็นข้อที่ 1 ดังนั้น

$$P(A') = 1 - P(A) = 1 - 0.3 = 0.7$$

(2) จากสมบัติของความน่าจะเป็นข้อที่ 5 ดังนั้น

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.2 - 0.1 = 0.4$$

(3) แทน A ด้วย $A \cup B$ ในสมบัติของความน่าจะเป็นข้อที่ 1 ดังนั้น

$$P[(A \cup B)'] = 1 - P(A \cup B) = 1 - 0.4 = 0.6$$

(4) เนื่องจาก $(A \cup B)' \cap B = \varnothing$ ดังนั้นเหตุการณ์ $(A \cup B)'$ และ B เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน โดยสัจพจน์ของความน่าจะเป็นข้อที่ (3) จะได้ว่า

$$P[(A \cup B)' \cup B] = P[(A \cup B)'] + P(B) = 0.6 + 0.2 = 0.8$$

(5) จากสมบัติของความน่าจะเป็นข้อที่ 4 ดังนั้น

$$P(A - B) = P(A) - P(A \cap B) = 0.3 - 0.1 = 0.2$$

(6) เนื่องจาก $A' \cap B = B \cap A' = B - A$ จะได้ว่า

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

จากสมบัติของความน่าจะเป็นข้อที่ 4 ดังนั้น

$$P(A' \cap B) = P(B - A) = P(B) - P(B \cap A) = 0.2 - 0.1 = 0.1$$

(7) จากสมบัติของความน่าจะเป็นข้อที่ 5 พิจารณา

$$P(A' \cup B) = P(A') + P(B) - P(A' \cap B) = 0.7 + 0.2 - 0.1 = 0.8$$

■

<u>ตัวอย่างที่ 1.4.5</u> กำหนดให้ A, B และ C เป็นเหตุการณ์ในปริภูมิตัวอย่างที่ไม่เกิดร่วมกัน โดยที่ $A \cup B \cup C = S$, $P(B) = 0.1$ และ $P(C) = 0.5$ จงหา $P(A)$

วิธีทำ เนื่องจาก A, B และ C เป็นเหตุการณ์ในปริภูมิตัวอย่างที่ไม่เกิดร่วมกัน โดยที่ $A \cup B \cup C = S$ และจากสมบัติของความน่าจะเป็นข้อที่ 7 ดังนั้น

$$P(A) + P(B) + P(C) = 1$$

นั่นคือ

$$P(A) = 1 - (P(B) + P(C)) = 1 - (0.1 + 0.5) = 0.4$$

ดังนั้น ความน่าจะเป็นของเหตุการณ์ A มีค่าเท่ากับ 0.4

■

<u>ตัวอย่างที่ 1.4.6</u> ประธานค่ายอาสาฯ ค่ายหนึ่ง ต้องการรับสมัครคนเพื่อจัดค่ายจำนวน 7 คน แต่มีผู้สมัครจำนวน 11 คน เป็นชาย 6 คน และเป็นหญิง 5 คน ถ้าเลือกคนเพื่อจัดค่ายโดยการสุ่ม จงหาความน่าจะเป็นที่เลือกคนเพื่อมาจัดค่ายได้ดังนี้

(1) เป็นชายอย่างน้อย 4 คน

(2) เป็นหญิงอย่างมาก 4 คน

วิธีทำ (1) ให้ A_i แทนเหตุการณ์ที่เลือกคนเพื่อจัดค่ายเป็นชาย i คน โดยที่ $i = 2,3,4,5,6$ ซึ่งเห็นได้ชัดเจนว่า $A_2, \ldots, A_6$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน เหตุการณ์ที่เลือกคนเพื่อจัดค่ายเป็นชายอย่างน้อย 4 คน สามารถแทนได้ด้วยเหตุการณ์ $A_4 \cup A_5 \cup A_6$ เนื่องจากเหตุการณ์ A_4, A_5 และ A_6 เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน โดยสัจพจน์ของความน่าจะเป็นข้อที่ (3) จะได้ว่า

$$\begin{aligned}
P(A_4 \cup A_5 \cup A_6) &= P(A_4) + P(A_5) + P(A_6) \\
&= \frac{{}^6C_4\,{}^5C_3}{{}^{11}C_7} + \frac{{}^6C_5\,{}^5C_2}{{}^{11}C_7} + \frac{{}^6C_6\,{}^5C_1}{{}^{11}C_7} \\
&= \frac{150}{330} + \frac{60}{330} + \frac{5}{330} \\
&= \frac{215}{330} = 0.6515
\end{aligned}$$

ดังนั้น ความน่าจะเป็นที่เลือกคนเพื่อจัดค่ายเป็นชายอย่างน้อย 4 คน มีค่าเท่ากับ 0.6515

(2) ให้ B_i แทนเหตุการณ์ที่เลือกคนเข้ามาทำงานเป็นหญิง i คน โดยที่ $i = 1,2,3,4,5$ ซึ่งเห็นได้ชัดเจนว่า $B_1,\dots,B_5$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน เหตุการณ์ที่เลือกคนเข้ามาทำงานเป็นหญิงอย่างมาก 4 คน สามารถแทนได้ด้วยเหตุการณ์ $B_1 \cup B_2 \cup B_3 \cup B_4$ หรือ B_5' จากสมบัติของความน่าจะเป็นข้อที่ 1 จึงได้ว่า

$$P(B_5') = 1 - P(B_5) = 1 - \frac{{}^{6}C_2\,{}^{5}C_5}{{}^{11}C_7} = 1 - \frac{15}{330} = \frac{315}{330} = 0.9545$$

ดังนั้น ความน่าจะเป็นที่เลือกคนเพื่อจัดค่ายเป็นหญิงอย่างมาก 4 คน มีค่าเท่ากับ 0.9545 ■

<u>ตัวอย่างที่ 1.4.7</u> การศึกษาความสัมพันธ์ระหว่างประเภทของหอพัก ชั้นปี และเพศของนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ของมหาวิทยาลัยแห่งหนึ่งซึ่งมีจำนวน 160 คน ได้ข้อมูลจำแนกตามเพศ ชั้นปี และประเภทของหอพัก ดังนี้

ชั้นปีที่	หอพักในมหาวิทยาลัย		หอพักนอกมหาวิทยาลัย	
	ชาย	หญิง	ชาย	หญิง
1	12	12	1	1
2	10	7	20	17
3	9	8	15	12
4	5	2	11	10
อื่น ๆ	2	1	5	0
รวม	38	30	52	40

ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน จงหาความน่าจะเป็นที่

(1) นักศึกษาที่สุ่มมาไม่ใช่นักศึกษาชั้นปีที่ 3

(2) นักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 3 หรือเป็นนักศึกษาชาย

(3) นักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 4 ซึ่งอยู่หอพักนอกมหาวิทยาลัย

วิธีทำ กำหนดให้ A แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 3

B แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาชาย

C แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 4

D แทนเหตุการณ์ที่นักศึกษาอยู่หอพักนอกมหาวิทยาลัย

(1) เหตุการณ์ที่นักศึกษาไม่ใช่นักศึกษาชั้นปีที่ 3 แทนได้ด้วย A' จากสมบัติของความน่าจะเป็นข้อที่ 1 จะได้ว่า

$$P(A') = 1 - P(A) = 1 - \frac{44}{160} = \frac{116}{160} = 0.725$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน ความน่าจะเป็นที่นักศึกษาที่สุ่มมาไม่ใช่นักศึกษาชั้นปีที่ 3 มีค่าเท่ากับ 0.725

(2) เหตุการณ์ที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 3 หรือเป็นนักศึกษาชายแทนได้ด้วย $A \cup B$ จากสมบัติของความน่าจะเป็นข้อที่ 5 จะได้ว่า

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ &= \frac{44}{160} + \frac{90}{160} - \frac{24}{160} = \frac{110}{160} = 0.6875 \end{aligned}$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน ความน่าจะเป็นที่นักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 3 หรือเป็นนักศึกษาชายมีค่าเท่ากับ 0.6875

(3) เหตุการณ์ที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 4 ซึ่งอยู่หอพักนอกมหาวิทยาลัยแทนได้ด้วย $C \cap D$ พิจารณา

$$P(C \cap D) = \frac{21}{160} = 0.1313$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน ความน่าจะเป็นที่นักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 4 ซึ่งอยู่หอพักนอกมหาวิทยาลัยมีค่าเท่ากับ 0.1313 ■

1.5 ความน่าจะเป็นแบบมีเงื่อนไข

การศึกษาเรื่องความน่าจะเป็น บางครั้งเราจำเป็นต้องหาค่าความน่าจะเป็นของเหตุการณ์ที่มีเงื่อนไขว่าต้องมีเหตุการณ์อีกเหตุการณ์หนึ่งเกิดขึ้นก่อน จึงเป็นที่มาของความน่าจะเป็นแบบมีเงื่อนไข ซึ่งมีนิยามดังนี้

นิยามที่ 1.5.1 กำหนดให้ A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B \subset S$) โดยที่ $P(B) > 0$ ดังนั้น ***ความน่าจะเป็นแบบมีเงื่อนไข (conditional probability)*** ของเหตุการณ์ A เมื่อเหตุการณ์ B ได้เกิดขึ้นแล้ว เขียนแทนด้วยสัญลักษณ์ $P(A|B)$ ซึ่งมีนิยามดังนี้

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

<u>ข้อสังเกต</u> จากนิยามของความน่าจะเป็นแบบมีเงื่อนไข

1. สำหรับเหตุการณ์ A และ B ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B \subset S$) โดยที่ $P(B) > 0$ แล้ว

$$P(A \cap B) = P(B)P(A|B)$$

2. สำหรับเหตุการณ์ A และ B ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B \subset S$) โดยที่ $P(A) > 0$ แล้ว

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

หรือ

$$P(B \cap A) = P(A)P(B|A)$$

แต่เนื่องจาก $P(A \cap B) = P(B \cap A)$ ดังนั้น

$$P(A \cap B) = P(A)P(B \mid A)$$

3. สำหรับเหตุการณ์ A, B และ C ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B, C \subset S$) โดยที่ $P(A \cap B) > 0$ พิจารณา

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

ดังนั้น

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

4. สำหรับเหตุการณ์ A ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A \subset S$) แล้ว

$$P(A) = P(A \mid S)$$

<u>ข้อสังเกต</u>

ความน่าจะเป็นแบบมีเงื่อนไข สอดคล้องกับสัจพจน์ของความน่าจะเป็น และสมบัติของความน่าจะเป็นในหัวข้อที่ 1.4.3 และ 1.4.4

<u>ตัวอย่างที่ 1.5.1</u> กำหนดให้ A และ B เป็นเหตุการณ์ในปริภูมิตัวอย่าง โดยที่ $P(A) = 0.68$, $P(B) = 0.55$ และ $P(A \cap B) = 0.32$ จงหาความน่าจะเป็นต่อไปนี้

(1) $P(A \mid B)$

(2) $P(B \mid A)$

(3) $P(B' \mid A)$

(4) $P(B \mid A')$

(5) $P(B' \mid A')$

วิธีทำ (1) จากนิยามของความน่าจะเป็นแบบมีเงื่อนไข ดังนั้น

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.32}{0.55} = \frac{32}{55}$$

(2) จากนิยามของความน่าจะเป็นแบบมีเงื่อนไข ดังนั้น

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} = \frac{0.32}{0.68} = \frac{8}{17}$$

(3) จากนิยามของความน่าจะเป็นแบบมีเงื่อนไขจะได้ว่า

$$P(B' \mid A) = \frac{P(B' \cap A)}{P(A)}$$

และจากสมบัติของเซต $B' \cap A = A \cap B' = A - B$ นั่นคือ $P(B' \cap A) = P(A - B)$ ดังนั้น

$$P(B' \mid A) = \frac{P(A-B)}{P(A)}$$

จากสมบัติของความน่าจะเป็นข้อที่ 4 ในหัวข้อที่ 1.4.4 ทำให้ได้ว่า $P(A-B) = P(A) - P(A \cap B)$ ดังนั้น

$$P(B' \mid A) = \frac{P(A) - P(A \cap B)}{P(A)} = \frac{0.68 - 0.32}{0.68} = \frac{0.36}{0.68} = \frac{9}{17}$$

(4) จากนิยามของความน่าจะเป็นแบบมีเงื่อนไขจะได้ว่า

$$P(B \mid A') = \frac{P(B \cap A')}{P(A')}$$

และจากสมบัติของเซต $B \cap A' = B - A$ นั่นคือ $P(B \cap A') = P(B - A)$ ดังนั้น

$$P(B \mid A') = \frac{P(B-A)}{P(A')}$$

จากสมบัติของความน่าจะเป็นข้อที่ 1 และ 4 ในหัวข้อที่ 1.4.4 ทำให้ได้ว่า $P(A') = 1 - P(A)$ และ $P(B - A) = P(B) - P(B \cap A)$ ดังนั้น

$$P(B \mid A') = \frac{P(B) - P(B \cap A)}{1 - P(A)} = \frac{0.55 - 0.32}{1 - 0.68} = \frac{0.23}{0.32} = \frac{23}{32}$$

(5) จากนิยามของความน่าจะเป็นแบบมีเงื่อนไขจะได้ว่า

$$P(B' \mid A') = \frac{P(B' \cap A')}{P(A')}$$

และจากสมบัติของเซต $B' \cap A' = (B \cup A)'$ นั่นคือ $P(B' \cap A') = P((B \cup A)')$ ดังนั้น

$$P(B' \mid A') = \frac{P((B \cup A)')}{P(A')}$$

จากสมบัติของความน่าจะเป็นข้อที่ 1 ในหัวข้อที่ 1.4.4 ทำให้ได้ว่า $P(A') = 1 - P(A)$ และ $P((B \cup A)') = 1 - P(B \cup A)$ ดังนั้น

$$P(B' \mid A') = \frac{1 - P(B \cup A)}{1 - P(A)}$$

จากสมบัติของความน่าจะเป็นข้อที่ 5 ในหัวข้อที่ 1.4.4 และสมบัติของเซต ทำให้ได้ว่า $P(B \cup A) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ดังนั้น

$$P(B' \mid A') = \frac{1 - (P(A) + P(B) - P(A \cap B))}{1 - P(A)} = \frac{1 - (0.68 + 0.55 - 0.32)}{1 - 0.68} = \frac{0.09}{0.32} = \frac{9}{32}$$ ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 1.5.1 ข้อที่ (1) และ (2) จะเห็นได้ว่า $P(A|B) \neq P(B|A)$ ดังนั้น การคำนวณหาค่าความน่าจะเป็นแบบมีเงื่อนไข จำเป็นจะต้องพิจารณาให้ดีว่าเหตุการณ์ใดเกิดก่อน เหตุการณ์ใดเกิดหลัง หรือเหตุการณ์ใดเป็นเงื่อนไข

<u>ตัวอย่างที่ 1.5.2</u> การศึกษาความสัมพันธ์ระหว่างประเภทของหอพัก ชั้นปี และเพศของนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ของมหาวิทยาลัยแห่งหนึ่งซึ่งมีจำนวน 160 คน ได้ข้อมูลจำแนกตามเพศ ชั้นปี และประเภทของหอพัก ดังนี้

ชั้นปีที่	หอพักในมหาวิทยาลัย		หอพักนอกมหาวิทยาลัย	
	ชาย	หญิง	ชาย	หญิง
1	12	12	1	1
2	10	7	20	17
3	9	8	15	12
4	5	2	11	10
อื่น ๆ	2	1	5	0
รวม	38	30	52	40

ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน จงหาความน่าจะเป็นที่

(1) นักศึกษาเป็นนักศึกษาชั้นปีที่ 2 ถ้าทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชาย

(2) นักศึกษาเป็นนักศึกษาชาย ถ้าทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 1

(3) นักศึกษาเป็นนักศึกษาหญิง ถ้าทราบว่านักศึกษาที่สุ่มมาอยู่หอพักในมหาวิทยาลัย

วิธีทำ กำหนดให้ A แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 2

B แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาชาย

C แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 1

D แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาหญิง

E แทนเหตุการณ์ที่นักศึกษาอยู่หอพักในมหาวิทยาลัย

(1) เนื่องจากทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชาย เหตุการณ์นี้จึงเป็นเงื่อนไข ความน่าจะเป็นที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 2 ถ้าทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชาย แทนได้ด้วย $P(A|B)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{30}{160} \cdot \frac{160}{90} = \frac{30}{90} = 0.3333$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน ความน่าจะเป็นที่นักศึกษาเป็นนักศึกษาชั้นปีที่ 2 ถ้าทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชายมีค่าเท่ากับ 0.3333

(2) เนื่องจากทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 1 เหตุการณ์นี้จึงเป็นเงื่อนไข ความน่าจะเป็นที่นักศึกษาเป็นนักศึกษาชาย ถ้าทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 1 แทนได้ด้วย $P(B|C)$

$$P(B|C)=\frac{P(B\cap C)}{P(C)}=\frac{13}{160}\cdot\frac{160}{26}=\frac{13}{26}=0.5$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน ความน่าจะเป็นที่นักศึกษาเป็นนักศึกษาชาย ถ้าทราบว่านักศึกษาที่สุ่มมาเป็นนักศึกษาชั้นปีที่ 1 มีค่าเท่ากับ 0.5

(3) เนื่องจากทราบว่านักศึกษาที่สุ่มมาอยู่หอพักในมหาวิทยาลัย เหตุการณ์นี้จึงเป็นเงื่อนไข ความน่าจะเป็นที่นักศึกษาเป็นนักศึกษาหญิง ถ้าทราบว่านักศึกษาที่สุ่มมาอยู่หอพักในมหาวิทยาลัย แทนได้ด้วย $P(D|E)$

$$P(D|E)=\frac{P(D\cap E)}{P(E)}=\frac{30}{160}\cdot\frac{160}{68}=\frac{30}{68}=0.4412$$

ดังนั้น ถ้าสุ่มนักศึกษากลุ่มนี้มา 1 คน ความน่าจะเป็นที่นักศึกษาเป็นนักศึกษาหญิง ถ้าทราบว่านักศึกษาที่สุ่มมาอยู่หอพักในมหาวิทยาลัยมีค่าเท่ากับ 0.4412 ■

ตัวอย่างที่ 1.5.3 จากการสำรวจลูกค้าที่ซื้อกล้องถ่ายรูปดิจิตอลพบว่า 60% ของทั้งหมดจะซื้อการ์ดความจำเพิ่ม และ 40% ของทั้งหมดจะซื้อแบตเตอรีเพิ่ม ส่วน 30% ของทั้งหมดจะซื้อทั้งการ์ดความจำและแบตเตอรีเพิ่ม จงหา

(1) ความน่าจะเป็นที่ผู้ซื้อแบตเตอรีเพิ่มจะซื้อหน่วยความจำเพิ่มด้วย

(2) ความน่าจะเป็นที่ผู้ซื้อการ์ดความจำเพิ่มจะซื้อแบตเตอรีเพิ่มด้วย

วิธีทำ กำหนดให้ B แทนเหตุการณ์ที่ลูกค้าจะซื้อแบตเตอรีเพิ่ม

C แทนเหตุการณ์ที่ลูกค้าจะซื้อการ์ดความจำเพิ่ม

จากโจทย์ 60% ของทั้งหมดจะซื้อการ์ดความจำเพิ่ม และ 40% ของทั้งหมดจะซื้อแบตเตอรีเพิ่ม ส่วน 30% ของทั้งหมดจะซื้อทั้งการ์ดความจำและแบตเตอรีเพิ่ม ดังนั้น

$$P(C)=0.6\quad P(B)=0.4 \text{ และ } P(B\cap C)=0.3$$

(1) พิจารณาลูกค้าที่ซื้อแบตเตอรีเพิ่มเป็นเงื่อนไข ความน่าจะเป็นที่ผู้ซื้อแบตเตอรีเพิ่มจะซื้อหน่วยความจำเพิ่มด้วย แทนด้วยสัญลักษณ์ $P(C|B)$ ซึ่งหาค่าได้ตามนิยามของความน่าจะเป็นแบบมีเงื่อนไข ดังนี้

$$P(C|B)=\frac{P(C\cap B)}{P(B)}=\frac{0.3}{0.4}=0.75$$

ดังนั้น ความน่าจะเป็นที่ผู้ซื้อแบตเตอรีเพิ่มจะซื้อการ์ดความจำเพิ่มด้วยมีค่าเท่ากับ 0.75

(2) พิจารณาลูกค้าที่ซื้อการ์ดความจำเพิ่มเป็นเงื่อนไข ความน่าจะเป็นที่ผู้ซื้อการ์ดความจำเพิ่มจะซื้อแบตเตอรีเพิ่มด้วย แทนด้วยสัญลักษณ์ $P(B|C)$ ซึ่งหาค่าได้ตามนิยามของความน่าจะเป็นแบบมีเงื่อนไข ดังนี้

$$P(B|C)=\frac{P(B\cap C)}{P(C)}=\frac{0.3}{0.6}=0.5$$

ดังนั้น ความน่าจะเป็นที่ผู้ซื้อการ์ดความจำเพิ่มจะซื้อแบตเตอรีเพิ่มด้วยมีค่าเท่ากับ 0.5 ■

<u>**ตัวอย่างที่ 1.5.4**</u> จากการสำรวจนักศึกษาของมหาวิทยาลัยแห่งหนึ่ง พบว่าเป็นผู้ชาย 70% และเป็นผู้หญิง 30% และยังทราบว่า 40% ของนักศึกษาผู้ชาย และ 60% ของนักศึกษาผู้หญิงเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ถ้าทำการสุ่มเลือกนักศึกษามา 1 คน จงหาความน่าจะเป็นที่สุ่มเลือกได้นักศึกษาวิศวกรรมศาสตร์

วิธีทำ กำหนดให้ M แทนเหตุการณ์ที่นักศึกษาเป็นผู้ชาย ดังนั้น M' แทนเหตุการณ์ที่นักศึกษาเป็นผู้หญิง
E แทนเหตุการณ์ที่นักศึกษาเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์

จากโจทย์ นักศึกษาเป็นผู้ชาย 70% และเป็นผู้หญิง 30% ดังนั้น $P(M)=0.7$ และ $P(M')=0.3$

ส่วน 40% ของนักศึกษาผู้ชายเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ นั่นหมายความว่า เราพิจารณาเฉพาะนักศึกษาผู้ชายแล้วจึงพิจารณาต่อไปอีก พบว่าเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ 40% ดังนั้น

$$P(E\,|\,M)=0.4$$

และ 60% ของนักศึกษาผู้หญิงเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ นั่นหมายความว่า เราพิจารณาเฉพาะนักศึกษาผู้หญิงแล้วจึงพิจารณาต่อไปอีก พบว่าเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ 60% ดังนั้น

$$P(E\,|\,M')=0.6$$

ความน่าจะเป็นที่จะสุ่มเลือกได้นักศึกษาวิศวกรรมศาสตร์ แทนได้ด้วยสัญลักษณ์ $P(E)$ ซึ่งสามารถหาค่าได้โดยพิจารณา $E=E\cap S=E\cap(M\cup M')=(E\cap M)\cup(E\cap M')$ นั่นคือ

$$P(E)=P((E\cap M)\cup(E\cap M'))$$

เนื่องจาก $(E\cap M)\cap(E\cap M')=\varnothing$ นั่นคือ เหตุการณ์ $E\cap M$ และ $E\cap M'$ เป็นอิสระต่อกัน โดยใช้สัจพจน์ของความน่าจะเป็นข้อที่ 3 จะได้ว่า

$$P(E)=P(E\cap M)+P(E\cap M')$$

และจากหมายเหตุข้อที่ 1 หลังนิยามที่ 1.5.1 จะได้ว่า

$$P(E)=P(E\,|\,M)P(M)+P(E\,|\,M')P(M')=(0.4)(0.7)+(0.6)(0.3)=0.46$$

ดังนั้น ถ้าทำการสุ่มเลือกนักศึกษามา 1 คน ความน่าจะเป็นที่สุ่มเลือกได้นักศึกษาวิศวกรรมศาสตร์มีค่าเท่ากับ 0.46 ■

1.6 ทฤษฎีบทของเบส์ (Bayes' Theorem)

ในหัวข้อนี้จะกล่าวถึงทฤษฎีบทของเบส์ ซึ่งเป็นทฤษฎีบทสำคัญที่ช่วยให้การคำนวณหาค่าความน่าจะเป็นแบบมีเงื่อนไขมีความสะดวกกว่าการคำนวณโดยใช้สัจพจน์หรือสมบัติของความน่าจะเป็น

ทฤษฎีบทที่ 1.6.1 กำหนดให้ $C_1, C_2, \ldots, C_n$ เป็นเหตุการณ์ในปริภูมิตัวอย่างที่ไม่เกิดร่วมกัน (นั่นคือ $C_1, C_2, \ldots, C_n \subset S$ และ $C_i \cap C_j = \varnothing$,$\forall i \neq j$ โดยที่ $i, j = 1,2,\ldots,n$) และ

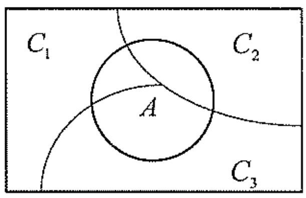
$$\bigcup_{i=1}^{n} C_i = S$$

ถ้า A เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A \subset S$) แล้ว

$$P(A) = P(C_1)P(A|C_1) + \ldots + P(C_n)P(A|C_n) = \sum_{i=1}^{n} P(C_i)P(A|C_i)$$

พิสูจน์ กำหนดให้ $C_1, C_2, \ldots, C_n$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน (นั่นคือ $C_i \cap C_j = \varnothing$,$\forall i \neq j$ โดยที่ $i, j = 1,2,\ldots,n$) และ $\bigcup_{i=1}^{n} C_i = S$ ดังนั้น สำหรับเหตุการณ์ A ใด ๆ ในปริภูมิตัวอย่าง (สามารถดูตัวอย่างแผนภาพเวนน์และออยเลอร์ในกรณีที่ $n = 3$ ได้ในรูปที่ 1.6.1) จะได้ว่า

$$A = (A \cap C_1) \cup (A \cap C_2) \cup \ldots \cup (A \cap C_n)$$



รูปที่ 1.6.1 ตัวอย่างแผนภาพเวนน์และออยเลอร์ในกรณีที่ $n = 3$

พิจารณา

$$P(A) = P((A \cap C_1) \cup (A \cap C_2) \cup \ldots \cup (A \cap C_n))$$

เนื่องจาก $(A \cap C_i) \subset C_i$ สำหรับทุก $i = 1,2,\ldots,n$ และจากที่กำหนดให้ $C_i \cap C_j = \varnothing$,$\forall i \neq j$ โดยที่ $i, j = 1,2,\ldots,n$ ดังนั้น สำหรับทุกค่า $i \neq j$ เมื่อ $i, j = 1,2,\ldots,n$

$$(A \cap C_i) \cap (A \cap C_j) = \varnothing$$

นั่นคือ $A \cap C_1, A \cap C_2, \ldots, A \cap C_n$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน ทำให้ได้ว่า

$$P((A \cap C_1) \cup (A \cap C_2) \cup \ldots \cup (A \cap C_n)) = P(A \cap C_1) + P(A \cap C_2) + \ldots + P(A \cap C_n)$$

ดังนั้น

$$P(A) = P(A \cap C_1) + P(A \cap C_2) + \ldots + P(A \cap C_n)$$

และจากหมายเหตุข้อที่ 1 ท้ายบทนิยามที่ 1.5.1 จะได้ว่า $P(A \cap C_i) = P(C_i)P(A \mid C_i)$ สำหรับทุก $i = 1, 2, \ldots, n$ ดังนั้น

$$P(A) = P(C_1)P(A \mid C_1) + \ldots + P(C_n)P(A \mid C_n) = \sum_{i=1}^{n} P(C_i)P(A \mid C_i)$$

<u>หมายเหตุ</u>

$P(A)$ ในทฤษฎีบทข้างต้นเรียกว่า ***ความน่าจะเป็นรวม (total probability)*** ของเหตุการณ์ A

<u>ทฤษฎีบทที่ 1.6.2</u> ทฤษฎีบทของเบส์ (Bayes' Theorem) กำหนดให้ $C_1, C_2, \ldots, C_n$ เป็นเหตุการณ์ในปริภูมิตัวอย่างที่ไม่เกิดร่วมกัน (นั่นคือ $C_1, C_2, \ldots, C_n \subset S$ และ $C_i \cap C_j = \varnothing$ $, \forall i \neq j$ โดยที่ $i, j = 1, 2, \ldots, n$) และ

$$\bigcup_{i=1}^{n} C_i = S$$

ถ้า A เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A \subset S$) และ $P(A) \neq 0$ แล้ว

$$P(C_j \mid A) = \frac{P(C_j)P(A \mid C_j)}{\sum_{i=1}^{n} P(C_i)P(A \mid C_i)}, \; j = 1, 2, \ldots, n$$

<u>พิสูจน์</u> กำหนดให้ $C_1, C_2, \ldots, C_n$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน (นั่นคือ $C_i \cap C_j = \varnothing$ $, \forall i \neq j$ โดยที่ $i, j = 1, 2, \ldots, n$) และ $\bigcup_{i=1}^{n} C_i = S$ จากทฤษฎีบทที่ 1.6.1 ดังนั้น สำหรับเหตุการณ์ A ใด ๆ ในปริภูมิตัวอย่าง จะได้ว่า

$$P(A) = \sum_{i=1}^{n} P(C_i)P(A \mid C_i)$$

สำหรับทุก $j = 1, 2, \ldots, n$ โดยนิยามของความน่าจะเป็นแบบมีเงื่อนไข และ $P(A) \neq 0$ จะได้ว่า

$$P(C_j \mid A) = \frac{P(C_j \cap A)}{P(A)}$$

จากหมายเหตุข้อที่ 1 ท้ายบทนิยามที่ 1.5.1 จะได้ว่า $P(A \cap C_j) = P(C_j)P(A \mid C_j)$ สำหรับทุก $j = 1, 2, \ldots, n$ ดังนั้น

$$P(C_j \mid A) = \frac{P(C_j)P(A \mid C_j)}{P(A)}$$

แต่จาก $P(A) = \sum_{i=1}^{n} P(C_i)P(A \mid C_i)$ ดังนั้น สำหรับทุก $j = 1, 2, \ldots, n$

$$P(C_j \mid A) = \frac{P(C_j)P(A \mid C_j)}{\sum_{i=1}^{n} P(C_i)P(A \mid C_i)}$$

<u>หมายเหตุ</u>

$P(C_j)$ และ $P(C_j \mid A)$ ในทฤษฎีบทของเบส์ เรียกว่า ***ความน่าจะเป็นก่อน (prior probability)*** และ ***ความน่าจะเป็นภายหลัง (posterior probability)*** ตามลำดับ

<u>ตัวอย่างที่ 1.6.1</u> บริษัทแห่งหนึ่งทำการเก็บรวบรวมข้อมูลเกี่ยวกับประเภทของมหาวิทยาลัยที่พนักงานจบการศึกษา และระยะเวลาในการเลื่อนตำแหน่งของพนักงานพบว่า 80% ของพนักงานทั้งหมดจบการศึกษาจากมหาวิทยาลัยของรัฐบาล ส่วนที่เหลือจบการศึกษาจากมหาวิทยาลัยของเอกชน และ 40% ของพนักงานที่จบการศึกษาจากมหาวิทยาลัยของรัฐบาลจะได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี และ 15% ของพนักงานที่จบการศึกษาจากมหาวิทยาลัยของเอกชนจะได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี ถ้ามีพนักงานคนหนึ่งที่ได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี จงหาความน่าจะเป็นที่

(1) พนักงานคนนี้จบการศึกษาจากมหาวิทยาลัยของรัฐบาล

(2) พนักงานคนนี้จบการศึกษาจากมหาวิทยาลัยของเอกชน

วิธีทำ กำหนดให้ C_1 แทนเหตุการณ์ที่พนักงานจบการศึกษาจากมหาวิทยาลัยของรัฐบาล

C_2 แทนเหตุการณ์ที่พนักงานจบการศึกษาจากมหาวิทยาลัยของเอกชน

A แทนเหตุการณ์ที่พนักงานได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี

<u>สิ่งที่ต้องการหา</u>

(1) ถ้ามีพนักงานคนหนึ่งที่ได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี ความน่าจะเป็นที่พนักงานคนนี้จบการศึกษาจากมหาวิทยาลัยของรัฐบาล แทนได้ด้วยสัญลักษณ์ $P(C_1 \mid A)$

(2) ถ้ามีพนักงานคนหนึ่งที่ได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี ความน่าจะเป็นที่พนักงานคนนี้จบการศึกษาจากมหาวิทยาลัยของเอกชน แทนได้ด้วยสัญลักษณ์ $P(C_2 \mid A)$

<u>สิ่งที่โจทย์กำหนดให้</u>

1. พบว่า 80% ของพนักงานทั้งหมดจบการศึกษาจากมหาวิทยาลัยของรัฐบาล ส่วนที่เหลือจบการศึกษาจากมหาวิทยาลัยของเอกชน ดังนั้น $P(C_1) = 0.8$ และ $P(C_2) = 0.2$

2. พบว่า 40% ของพนักงานที่จบการศึกษาจากมหาวิทยาลัยของรัฐบาลจะได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี ดังนั้น $P(A \mid C_1) = 0.4$

3. พบว่า 15% ของพนักงานที่จบการศึกษาจากมหาวิทยาลัยของเอกชนจะได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี ดังนั้น $P(A \mid C_2) = 0.15$

<u>ตรวจสอบเงื่อนไขในการใช้ทฤษฎีบทของเบส์</u>

1. เนื่องจากปริภูมิตัวอย่างในตัวอย่างนี้คือพนักงานทั้งหมด ดังนั้น $C_1 \cup C_2 = S$

2. เนื่องจากพนักงานจะจบการศึกษาจากมหาวิทยาลัยของรัฐบาล หรือมหาวิทยาลัยของเอกชนที่ใดที่หนึ่งเท่านั้น ดังนั้น $C_1 \cap C_2 = \varnothing$

3. $P(A) = P(C_1)P(A|C_1) + P(C_2)P(A|C_2) = (0.8)(0.4) + (0.2)(0.15) = 0.35 \neq 0$

จากทฤษฎีบทของเบส์จะได้ว่า

(1) $P(C_1|A) = \dfrac{P(C_1)P(A|C_1)}{P(A)} = \dfrac{(0.8)(0.4)}{0.35} = 0.9143$

(2) $P(C_2|A) = \dfrac{P(C_2)P(A|C_2)}{P(A)} = \dfrac{(0.2)(0.15)}{0.35} = 0.0857$

ดังนั้น ถ้ามีพนักงานคนหนึ่งที่ได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี ความน่าจะเป็นที่พนักงานคนนี้จบการศึกษาจากมหาวิทยาลัยของรัฐบาล มีค่าเท่ากับ 0.9143 และถ้ามีพนักงานคนหนึ่งที่ได้รับการเลื่อนตำแหน่งภายในระยะเวลา 5 ปี ความน่าจะเป็นที่พนักงานคนนี้จบการศึกษาจากมหาวิทยาลัยของเอกชน มีค่าเท่ากับ 0.0857 ■

<u>ตัวอย่างที่ 1.6.2</u> การผลิตสินค้าชนิดหนึ่งสามารถผลิตจากเครื่องจักรเครื่องใดเครื่องหนึ่งซึ่งประกอบไปด้วยเครื่องจักรเครื่องที่ 1, 2 และ 3 จากข้อมูลทราบว่าสินค้าที่ผลิตจากเครื่องจักรเครื่องที่ 1, 2 และ 3 จะถูกแบ่งไปเป็นสินค้าตัวอย่างคิดเป็น 7%, 9% และ 8% ตามลำดับ ถ้าสุ่มสินค้าตัวอย่างมา 1 ชิ้น จงหาความน่าจะเป็นที่

(1) สินค้าชิ้นนี้ผลิตจากเครื่องจักรเครื่องที่ 1

(2) สินค้าชิ้นนี้ผลิตจากเครื่องจักรเครื่องที่ 2

(3) สินค้าชิ้นนี้ผลิตจากเครื่องจักรเครื่องที่ 3

วิธีทำ กำหนดให้ C_1 แทนเหตุการณ์ที่สินค้าผลิตด้วยเครื่องจักรเครื่องที่ 1

C_2 แทนเหตุการณ์ที่สินค้าผลิตด้วยเครื่องจักรเครื่องที่ 2

C_3 แทนเหตุการณ์ที่สินค้าผลิตด้วยเครื่องจักรเครื่องที่ 3

A แทนเหตุการณ์ที่เป็นสินค้าตัวอย่าง

<u>สิ่งที่ต้องการหา</u>

(1) ถ้าสุ่มสินค้าตัวอย่างมา 1 ชิ้น ความน่าจะเป็นที่สินค้าชิ้นนี้ผลิตจากเครื่องจักรเครื่องที่ 1 แทนได้ด้วยสัญลักษณ์ $P(C_1|A)$

(2) ถ้าสุ่มสินค้าตัวอย่างมา 1 ชิ้น ความน่าจะเป็นที่สินค้าชิ้นนี้ผลิตจากเครื่องจักรเครื่องที่ 2 แทนได้ด้วยสัญลักษณ์ $P(C_2|A)$

(3) ถ้าสุ่มสินค้าตัวอย่างมา 1 ชิ้น ความน่าจะเป็นที่สินค้าชิ้นนี้ผลิตจากเครื่องจักรเครื่องที่ 3 แทนได้ด้วยสัญลักษณ์ $P(C_3|A)$

<u>สิ่งที่โจทย์กำหนดให้</u>

1. สินค้าที่ผลิตจากเครื่องจักรเครื่องที่ 1, 2 และ 3 จะถูกแบ่งไปเป็นสินค้าตัวอย่างคิดเป็น 7%, 9% และ 8% ตามลำดับ ดังนั้น $P(A|C_1) = 0.07$, $P(A|C_2) = 0.09$ และ $P(A|C_3) = 0.08$

2. เนื่องจากโจทย์ไม่ได้กำหนดว่าสินค้าผลิตจากเครื่องจักรเครื่องที่ 1, 2 และ 3 ด้วยสัดส่วนเท่าใด จึงสมมติว่าสินค้าผลิตจากเครื่องจักรทั้งสามเครื่องด้วยความน่าจะเป็นเท่า ๆ กัน ดังนั้น $P(C_1)=\frac{1}{3}$, $P(C_2)=\frac{1}{3}$ และ $P(C_3)=\frac{1}{3}$

<u>ตรวจสอบเงื่อนไขในการใช้ทฤษฎีบทของเบส์</u>

1. เนื่องจากปริภูมิตัวอย่างในตัวอย่างนี้คือสินค้าชนิดนี้ทั้งหมด ดังนั้น $C_1 \cup C_2 \cup C_3 = S$

2. เนื่องจากสินค้าแต่ละชิ้นจะผลิตจากเครื่องจักรเครื่องที่ 1, 2 หรือ 3 อย่างใดอย่างหนึ่งเท่านั้น ดังนั้น $C_1 \cap C_2 = C_1 \cap C_3 = C_2 \cap C_3 = \varnothing$

3. $P(A) = P(C_1)P(A|C_1) + P(C_2)P(A|C_2) + P(C_3)P(A|C_3)$

$$= \frac{1}{3}(0.07) + \frac{1}{3}(0.09) + \frac{1}{3}(0.08) = 0.08 \neq 0$$

จากทฤษฎีบทของเบส์จะได้ว่า

(1) $P(C_1|A) = \dfrac{P(C_1)P(A|C_1)}{P(A)} = \dfrac{\frac{1}{3}(0.07)}{0.08} = 0.2917$

(2) $P(C_2|A) = \dfrac{P(C_2)P(A|C_2)}{P(A)} = \dfrac{\frac{1}{3}(0.09)}{0.08} = 0.375$

(3) $P(C_3|A) = \dfrac{P(C_3)P(A|C_3)}{P(A)} = \dfrac{\frac{1}{3}(0.08)}{0.08} = 0.3333$

ดังนั้น ถ้าสุ่มสินค้าตัวอย่างมา 1 ชิ้น ความน่าจะเป็นที่สินค้าชิ้นนี้ผลิตจากเครื่องจักรเครื่องที่ 1, 2 และ 3 มีค่าเท่ากับ 0.2917, 0.375 และ 0.3333 ตามลำดับ ■

<u>ตัวอย่างที่ 1.6.3</u> จากการสำรวจเพศและสำนักวิชาที่สังกัดของนักศึกษามหาวิทยาลัยแห่งหนึ่ง พบว่าจากนักศึกษาทั้งหมดเป็นนักศึกษาชาย 70% นักศึกษาหญิง 30% และ 40% ของนักศึกษาชายเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ และ 60% ของนักศึกษาหญิงเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ถ้าทำการสุ่มเลือกนักศึกษามา 1 คน พบว่าเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ จงหาความน่าจะเป็นที่

(1) นักศึกษาที่สุ่มมาเป็นนักศึกษาชาย

(2) นักศึกษาที่สุ่มมาเป็นนักศึกษาหญิง

วิธีทำ กำหนดให้ C_1 แทนเหตุการณ์ที่สุ่มเลือกได้นักศึกษาชาย

C_2 แทนเหตุการณ์ที่สุ่มเลือกได้นักศึกษาหญิง

A แทนเหตุการณ์ที่สุ่มเลือกได้นักศึกษาที่สังกัดสำนักวิชาวิศวกรรมศาสตร์

<u>สิ่งที่ต้องการหา</u>

(1) ถ้าทำการสุ่มเลือกนักศึกษามา 1 คน พบว่าเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ความน่าจะเป็นที่นักศึกษาที่สุ่มมาเป็นนักศึกษาชาย แทนได้ด้วยสัญลักษณ์ $P(C_1 \mid A)$

(2) ถ้าทำการสุ่มเลือกนักศึกษามา 1 คน พบว่าเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ความน่าจะเป็นที่นักศึกษาที่สุ่มมาเป็นนักศึกษาหญิง แทนได้ด้วยสัญลักษณ์ $P(C_2 \mid A)$

<u>สิ่งที่โจทย์กำหนดให้</u>

1. จากนักศึกษาทั้งหมดเป็นนักศึกษาชาย 70% นักศึกษาหญิง 30% ดังนั้น $P(C_1) = 0.7$ และ $P(C_2) = 0.3$

2. พบว่า 40% ของนักศึกษาชายเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ดังนั้น $P(A \mid C_1) = 0.4$

3. พบว่า 60% ของนักศึกษาหญิงเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ดังนั้น $P(A \mid C_2) = 0.6$

<u>ตรวจสอบเงื่อนไขในการใช้ทฤษฎีบทของเบส์</u>

1. เนื่องจากปริภูมิตัวอย่างในตัวอย่างนี้คือนักศึกษาทั้งหมดของมหาวิทยาลัยแห่งนี้ ดังนั้น $C_1 \cup C_2 = S$

2. เนื่องจากนักศึกษาแต่ละคนจะเป็นนักศึกษาชายหรือนักศึกษาหญิงอย่างใดอย่างหนึ่งเท่านั้น ดังนั้น $C_1 \cap C_2 = \varnothing$

3. $P(A) = P(C_1)P(A \mid C_1) + P(C_2)P(A \mid C_2) = (0.7)(0.4) + (0.3)(0.6) = 0.46 \neq 0$

จากทฤษฎีบทของเบส์จะได้ว่า

(1) $P(C_1 \mid A) = \dfrac{P(C_1)P(A \mid C_1)}{P(A)} = \dfrac{(0.7)(0.4)}{0.46} = 0.6087$

(2) $P(C_2 \mid A) = \dfrac{P(C_2)P(A \mid C_2)}{P(A)} = \dfrac{(0.3)(0.6)}{0.46} = 0.3913$

ดังนั้น ถ้าทำการสุ่มเลือกนักศึกษามา 1 คน พบว่าเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ความน่าจะเป็นที่นักศึกษาที่สุ่มมาเป็นนักศึกษาชาย มีค่าเท่ากับ 0.6087 และถ้าทำการสุ่มเลือกนักศึกษามา 1 คน พบว่าเป็นนักศึกษาสำนักวิชาวิศวกรรมศาสตร์ ความน่าจะเป็นที่นักศึกษาที่สุ่มมาเป็นนักศึกษาหญิงมีค่าเท่ากับ 0. 3913 ■

<u>ตัวอย่างที่ 1.6.4</u> บ้านจัดสรรโครงการหนึ่งมีบ้านอยู่ 2 แบบ คือ บ้านชั้นเดียว และบ้านสองชั้น จากการสำรวจลูกค้าทั้งหมดที่เข้าไปเยี่ยมชมบ้านพบว่า 22% เลือกจองบ้านชั้นเดียว 42% เลือกจองบ้านสองชั้น ส่วนที่เหลือไม่สนใจจองบ้านของโครงการนี้ และทราบอีกว่าลูกค้าที่เลือกจองบ้านชั้นเดียวเป็นโสดถึง 44% ลูกค้าที่เลือกจองบ้านสองชั้นเป็นโสดถึง 21% และลูกค้าที่ไม่จองบ้านเป็นโสดถึง 35%

(1) ถ้าทำการสุ่มลูกค้าที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้ที่เป็นโสดมา 1 คน จงหาความน่าจะเป็นที่ลูกค้าคนนี้จะเลือกจองบ้านสองชั้น

(2) ถ้าทำการสุ่มลูกค้าที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้ที่ไม่โสดมา 1 คน จงหาความน่าจะเป็นที่ลูกค้าคนนี้จะเลือกจองบ้านชั้นเดียว

วิธีทำ กำหนดให้ C_1 แทนเหตุการณ์ที่ลูกค้าเลือกจองบ้านชั้นเดียว

C_2 แทนเหตุการณ์ที่ลูกค้าเลือกจองบ้านสองชั้น

C_3 แทนเหตุการณ์ที่ลูกค้าไม่จองบ้านเลย

A แทนเหตุการณ์ที่ลูกค้าเป็นโสด ดังนั้น A' แทนเหตุการณ์ที่ลูกค้าไม่โสด

<u>สิ่งที่ต้องการหา</u>

(1) ถ้าทำการสุ่มลูกค้าที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้ที่เป็นโสดมา 1 คน ความน่าจะเป็นที่ลูกค้าคนนี้จะเลือกจองบ้านสองชั้น แทนได้ด้วยสัญลักษณ์ $P(C_2 \mid A)$

(2) ถ้าทำการสุ่มลูกค้าที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้ที่ไม่โสดมา 1 คน จงหาความน่าจะเป็นที่ลูกค้าคนนี้จะเลือกจองบ้านชั้นเดียว แทนได้ด้วยสัญลักษณ์ $P(C_1 \mid A')$

<u>สิ่งที่โจทย์กำหนดให้</u>

1. จากการสำรวจลูกค้าทั้งหมดที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้พบว่า 22% เลือกจองบ้านชั้นเดียว 42% เลือกจองบ้านสองชั้น ส่วนที่เหลือไม่สนใจจองบ้านของโครงการนี้ ดังนั้น $P(C_1) = 0.22$, $P(C_2) = 0.42$ และ $P(C_3) = 1 - 0.22 - 0.42 = 0.36$

2. ลูกค้าที่เลือกจองบ้านชั้นเดียวเป็นโสดถึง 44% ดังนั้น $P(A \mid C_1) = 0.44$

3. ลูกค้าที่เลือกจองบ้านสองชั้นเป็นโสดถึง 21% ดังนั้น $P(A \mid C_2) = 0.21$

4. ลูกค้าที่ไม่จองบ้านเป็นโสดถึง 35% ดังนั้น $P(A \mid C_3) = 0.35$

<u>ตรวจสอบเงื่อนไขในการใช้ทฤษฎีบทของเบส์</u>

1. เนื่องจากปริภูมิตัวอย่างในตัวอย่างนี้คือลูกค้าทั้งหมดที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้ ดังนั้น $C_1 \cup C_2 \cup C_3 = S$

2. เนื่องจากลูกค้าที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้จะเลือกจองบ้านชั้นเดียว หรือจองบ้านสองชั้น หรือไม่จองบ้านเลยเพียงอย่างใดอย่างหนึ่งเท่านั้น ดังนั้น $C_1 \cap C_2 = C_1 \cap C_3 = C_2 \cap C_3 = \varnothing$

3. $P(A) = P(C_1)P(A \mid C_1) + P(C_2)P(A \mid C_2) + P(C_3)P(A \mid C_3)$

$$= (0.22)(0.44) + (0.42)(0.21) + (0.36)(0.35) = 0.311 \neq 0$$

ดังนั้น $P(A') = 1 - P(A) = 1 - 0.311 = 0.689 \neq 0$

จากทฤษฎีบทของเบส์จะได้ว่า

(1) $P(C_2 \mid A) = \dfrac{P(C_2)P(A \mid C_2)}{P(A)} = \dfrac{(0.42)(0.21)}{0.311} = 0.2836$

(2) $P(C_1 \mid A') = \dfrac{P(C_1)P(A' \mid C_1)}{P(A')} = \dfrac{P(C_1)(1 - P(A \mid C_1))}{P(A')} = \dfrac{(0.22)(1 - 0.44)}{0.689} = 0.1788$

ดังนั้น ถ้าทำการสุ่มลูกค้าที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้ที่เป็นโสดมา 1 คน ความน่าจะเป็นที่ลูกค้าคนนี้จะเลือกจองบ้านสองชั้นมีค่าเท่ากับ 0..2836 และถ้าทำการสุ่มลูกค้าที่เข้าไปเยี่ยมชมบ้านจัดสรรโครงการนี้ที่ไม่โสดมา 1 คน ความน่าจะเป็นที่ลูกค้าคนนี้จะเลือกจองบ้านชั้นเดียวมีค่าเท่ากับ 0.1788 ■

1.7 เหตุการณ์ที่เป็นอิสระต่อกัน

การศึกษาเรื่องความน่าจะเป็นของเหตุการณ์หลายเหตุการณ์ บางเหตุการณ์อาจมีผลกระทบต่อเหตุการณ์อื่น แต่บางเหตุการณ์อาจไม่มีผลกระทบต่อเหตุการณ์อื่น ซึ่งมีผลทำให้ความน่าจะเป็นแตกต่างกัน ดังนั้น ในหัวข้อนี้จะกล่าวถึงเหตุการณ์ที่เป็นอิสระต่อกัน

<u>นิยามที่ 1.7.1</u> กำหนดให้ A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B \subset S$) เหตุการณ์ A และ B จะเรียกว่าเป็น ***เหตุการณ์ที่เป็นอิสระต่อกัน (independent event)*** ก็ต่อเมื่อ การเกิดเหตุการณ์ A ไม่มีผลต่อการเกิดเหตุการณ์ B และการเกิดเหตุการณ์ B ไม่มีผลต่อการเกิดเหตุการณ์ A นั่นคือ

$$P(B \mid A) = P(B)$$
$$P(A \mid B) = P(A)$$

หรือ เหตุการณ์ A และ B เป็นอิสระต่อกัน ก็ต่อเมื่อ

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

<u>ข้อสังเกต</u>

1. กำหนดให้ A, B และ C เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B, C \subset S$) เหตุการณ์ A, B และ C เป็นเหตุการณ์ที่เป็นอิสระต่อกัน ก็ต่อเมื่อ

   (1) $P(A \cap B) = P(A)P(B)$

   (2) $P(A \cap C) = P(A)P(C)$

   (3) $P(B \cap C) = P(B)P(C)$

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

2. กำหนดให้ $A_1, A_2, \ldots, A_n$ เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A_1, A_2, \ldots, A_n \subset S$) ถ้า $A_1, A_2, \ldots, A_n$ เป็นเหตุการณ์ที่เป็นอิสระต่อกันแล้ว ทุก ๆ $2, 3, \ldots, n-1$ เหตุการณ์ใด ๆ ใน $A_1, A_2, \ldots, A_n$ จะต้องเป็นอิสระต่อกันด้วย

3. กำหนดให้ $A_1, A_2, \ldots, A_n$ เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A_1, A_2, \ldots, A_n \subset S$) ถ้า $A_1, A_2, \ldots, A_n$ เป็นเหตุการณ์ที่เป็นอิสระต่อกันแล้ว

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2)\ldots P(A_n)$$

<u>ตัวอย่างที่ 1.7.1</u> กำหนดให้ A, B และ C เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่างที่เป็นอิสระต่อกัน โดยที่ $P(A) = 0.2$, $P(B) = 0.3$ และ $P(C) = 0.4$ จงหาค่า

(1) $P(A \cap B)$

(2) $P(A \cap B \cap C)$

(3) $P(A \cup B \cup C)$

(4) $P[(A\cup B)\cap C]$

วิธีทำ (1) เนื่องจาก A, B และ C เป็นเหตุการณ์ที่เป็นอิสระต่อกัน ทำให้ได้ว่าเหตุการณ์ A และ B เป็นอิสระต่อกันด้วย ดังนั้น

$$P(A\cap B)=P(A)P(B)=(0.2)(0.3)=0.06$$

(2) เนื่องจาก A, B และ C เป็นเหตุการณ์ที่เป็นอิสระต่อกัน ดังนั้น

$$P(A\cap B\cap C)=P(A)P(B)P(C)=(0.2)(0.3)(0.4)=0.024$$

(3) จากสมบัติของความน่าจะเป็นข้อที่ 6 และ A, B และ C เป็นเหตุการณ์ที่เป็นอิสระต่อกัน ดังนั้น

$$\begin{aligned}P(A\cup B\cup C)&=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)\\&=P(A)+P(B)+P(C)-P(A)P(B)-P(A)P(C)-P(B)P(C)+P(A)P(B)P(C)\\&=0.2+0.3+0.4-(0.2)(0.3)-(0.2)(0.4)-(0.3)(0.4)+(0.2)(0.3)(0.4)=0.664\end{aligned}$$

(4) จากสมบัติของเซต สมบัติของความน่าจะเป็นข้อที่ 5 และ A, B และ C เป็นเหตุการณ์ที่เป็นอิสระต่อกัน ดังนั้น

$$\begin{aligned}P[(A\cup B)\cap C]&=P[(A\cap C)\cup(B\cap C)]\\&=P(A\cap C)+P(B\cap C)-P(A\cap B\cap C)\\&=P(A)P(C)+P(B)P(C)-P(A)P(B)P(C)\\&=(0.2)(0.4)+(0.3)(0.4)-(0.2)(0.3)(0.4)\\&=0.08+0.12-0.024=0.176\end{aligned}$$

■

ตัวอย่างที่ 1.7.2 กำหนดให้ A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง โดยที่ $P(A)=0.8$, $P(B)=0.5$ และ $P(A\cup B)=0.9$ จงพิจารณาว่าเหตุการณ์ต่อไปนี้เป็นอิสระต่อกันหรือไม่

(1) เหตุการณ์ A และ B

(2) เหตุการณ์ A' และ B

(3) เหตุการณ์ A และ B'

(4) เหตุการณ์ A' และ B'

วิธีทำ (1) การตรวจสอบว่าเหตุการณ์ A และ B เป็นอิสระต่อกันหรือไม่ พิจารณาได้ดังนี้

ถ้า $P(A\cap B)=P(A)P(B)$ แสดงว่า A และ B เป็นอิสระต่อกัน แต่ถ้า $P(A\cap B)\neq P(A)P(B)$ แสดงว่า A และ B ไม่เป็นอิสระต่อกัน พิจารณา

$$\begin{aligned}P(A\cup B)&=P(A)+P(B)-P(A\cap B)\\0.9&=0.8+0.5-P(A\cap B)\end{aligned}$$

นั่นคือ

$$P(A\cap B)=0.4$$

ต่อมาพิจารณา

$$P(A)P(B) = (0.8)(0.5) = 0.4$$

จะเห็นว่า $P(A \cap B) = P(A)P(B)$ ดังนั้น จึงสรุปว่าเหตุการณ์ A และ B เป็นอิสระต่อกัน

(2) การตรวจสอบว่าเหตุการณ์ A' และ B เป็นอิสระต่อกันหรือไม่ พิจารณาได้ดังนี้

ถ้า $P(A' \cap B) = P(A')P(B)$ แสดงว่า A' และ B เป็นอิสระต่อกัน แต่ถ้า $P(A' \cap B) \neq P(A')P(B)$ แสดงว่า A' และ B ไม่เป็นอิสระต่อกัน พิจารณา

$$\begin{aligned} P(A' \cap B) &= P(B \cap A') \\ &= P(B - A) \\ &= P(B) - P(B \cap A) \\ &= 0.5 - 0.4 \\ &= 0.1 \end{aligned}$$

ต่อมาพิจารณา

$$\begin{aligned} P(A')P(B) &= [1 - P(A)]P(B) \\ &= (1 - 0.8)(0.5) \\ &= 0.1 \end{aligned}$$

จะเห็นว่า $P(A' \cap B) = P(A')P(B)$ ดังนั้น จึงสรุปว่าเหตุการณ์ A' และ B เป็นอิสระต่อกัน

(3) การตรวจสอบว่าเหตุการณ์ A และ B' เป็นอิสระต่อกันหรือไม่ พิจารณาได้ดังนี้

ถ้า $P(A \cap B') = P(A)P(B')$ แสดงว่า A และ B' เป็นอิสระต่อกัน แต่ถ้า $P(A \cap B') \neq P(A)P(B')$ แสดงว่า A และ B' ไม่เป็นอิสระต่อกัน พิจารณา

$$\begin{aligned} P(A \cap B') &= P(A - B) \\ &= P(A) - P(A \cap B) \\ &= 0.8 - 0.4 \\ &= 0.4 \end{aligned}$$

ต่อมาพิจารณา

$$\begin{aligned} P(A)P(B') &= P(A)[1 - P(B)] \\ &= (0.8)(1 - 0.5) \\ &= 0.4 \end{aligned}$$

จะเห็นว่า $P(A \cap B') = P(A)P(B')$ ดังนั้น จึงสรุปว่าเหตุการณ์ A และ B' เป็นอิสระต่อกัน

(4) การตรวจสอบว่าเหตุการณ์ A' และ B' เป็นอิสระต่อกันหรือไม่ พิจารณาได้ดังนี้

ถ้า $P(A' \cap B') = P(A')P(B')$ แสดงว่า A' และ B' เป็นอิสระต่อกัน แต่ถ้า $P(A' \cap B') \neq P(A')P(B')$ แสดงว่า A' และ B' ไม่เป็นอิสระต่อกัน พิจารณา

$$P(A' \cap B') = P[(A \cup B)']$$

$$= 1 - P(A \cup B)$$
$$= 1 - 0.9$$
$$= 0.1$$

ต่อมาพิจารณา

$$P(A')P(B') = [1 - P(A)][1 - P(B)]$$
$$= (1 - 0.8)(1 - 0.5)$$
$$= 0.1$$

จะเห็นว่า $P(A' \cap B') = P(A')P(B')$ ดังนั้น จึงสรุปว่าเหตุการณ์ A' และ B' เป็นอิสระต่อกัน ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 1.7.2 กำหนดให้ A และ B เป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง (นั่นคือ $A, B \subset S$) ที่เป็นอิสระต่อกัน พิจารณา

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

จากสมบัติของความน่าจะเป็นข้อที่ 4 นั่นคือ $P(B - A) = P(B) - P(B \cap A)$ จะได้ว่า

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

จากความเป็นอิสระต่อกันของเหตุการณ์ A และ B และสมบัติของความน่าจะเป็นข้อที่ 1 จะได้ว่า

$$P(A' \cap B) = P(B) - P(B)P(A) = (1 - P(A))P(B) = P(A')P(B)$$

นั่นคือ

$$P(A' \cap B) = P(A')P(B)$$

ดังนั้น เหตุการณ์ A' และ B เป็นอิสระต่อกัน ในทำนองเดียวกัน ถ้าเหตุการณ์ A และ B เป็นอิสระต่อกันแล้ว เราสามารถสรุปได้ว่า

(1) เหตุการณ์ A' และ B เป็นอิสระต่อกัน
(2) เหตุการณ์ A และ B' เป็นอิสระต่อกัน
(3) เหตุการณ์ A' และ B' เป็นอิสระต่อกัน

<u>ตัวอย่างที่ 1.7.3</u> บริษัทแห่งหนึ่งมีรถยนต์เพื่อใช้งานอยู่ 2 คัน ซึ่งรถยนต์ทั้งสองคันนี้ทำงานเป็นอิสระต่อกัน ถ้าความน่าจะเป็นที่รถยนต์แต่ละคันพร้อมใช้งานมีค่าเท่ากับ 0.85 จงหา

(1) ความน่าจะเป็นที่จะมีรถยนต์พร้อมใช้งานทั้ง 2 คัน
(2) ความน่าจะเป็นที่จะมีรถยนต์พร้อมใช้งานอย่างน้อย 1 คัน

วิธีทำ กำหนดให้ W_1 แทนเหตุการณ์ที่รถยนต์คันที่ 1 พร้อมใช้งาน
W_2 แทนเหตุการณ์ที่รถยนต์คันที่ 2 พร้อมใช้งาน

จากโจทย์กำหนดว่าความน่าจะเป็นที่รถยนต์แต่ละคันพร้อมใช้งานมีค่าเท่ากับ 0.85 จะได้ว่า

$$P(W_1) = P(W_2) = 0.85$$

(1) เหตุการณ์ที่จะมีรถยนต์พร้อมใช้งานทั้ง 2 คัน หมายความว่า รถยนต์คันที่ 1 และคันที่ 2 พร้อมใช้งานทั้งคู่ ความน่าจะเป็นที่รถยนต์พร้อมใช้งานทั้ง 2 คัน แทนได้ด้วยสัญลักษณ์ $P(W_1 \cap W_2)$ และเนื่องจากรถยนต์ทั้งสองคันนี้ทำงานเป็นอิสระต่อกัน นั่นคือ W_1 และ W_2 เป็นอิสระต่อกัน ดังนั้น

$$P(W_1 \cap W_2) = P(W_1)P(W_2) = (0.85)(0.85) = 0.7225$$

ดังนั้น ความน่าจะเป็นที่จะมีรถยนต์พร้อมใช้งานทั้ง 2 คัน มีค่าเท่ากับ 0.7225

(2) เหตุการณ์ที่จะมีรถยนต์พร้อมใช้งานอย่างน้อย 1 คัน หมายความว่า รถยนต์คันที่ 1 พร้อมใช้งาน หรือรถยนต์คันที่ 2 พร้อมใช้งาน ความน่าจะเป็นที่รถยนต์พร้อมใช้งานแทนได้ด้วยสัญลักษณ์ $P(W_1 \cup W_2)$ และจากสมบัติของความน่าจะเป็นข้อที่ 5 ดังนั้น

$$P(W_1 \cup W_2) = P(W_1) + P(W_2) - P(W_1 \cap W_2)$$

เนื่องจากรถยนต์ทั้งสองคันนี้ทำงานเป็นอิสระต่อกัน นั่นคือ W_1 และ W_2 เป็นอิสระต่อกัน ดังนั้น

$$\begin{aligned} P(W_1 \cup W_2) &= P(W_1) + P(W_2) - P(W_1)P(W_2) \\ &= 0.85 + 0.85 - (0.85)(0.85) \\ &= 0.9775 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่จะมีรถยนต์พร้อมใช้งานอย่างน้อย 1 คัน มีค่าเท่ากับ 0.9775 ■

<u>ตัวอย่างที่ 1.7.4</u> พิจารณาการโยนเหรียญ 1 อัน 3 ครั้ง สมมติว่าความน่าจะเป็นที่เหรียญจะออกหัวในแต่ละครั้งมีค่าเท่ากับ 0.55 จงหา

(1) ความน่าจะเป็นที่เหรียญจะออกหัว 1 ครั้ง

(2) ความน่าจะเป็นที่เหรียญจะออกหัว 2 ครั้ง

(3) ความน่าจะเป็นที่เหรียญจะออกหัวทั้ง 3 ครั้ง

วิธีทำ กำหนดให้ H_i แทนเหตุการณ์ที่โยนได้หัวในครั้งที่ i เมื่อ $i = 1,2,3$

T_i แทนเหตุการณ์ที่โยนได้ก้อยในครั้งที่ i เมื่อ $i = 1,2,3$

จากโจทย์กำหนดความน่าจะเป็นที่เหรียญจะออกหัวในแต่ละครั้งมีค่าเท่ากับ 0.55 จะได้ว่า

$$P(H_1) = P(H_2) = P(H_3) = 0.55 \text{ และ } P(T_1) = P(T_2) = P(T_3) = 0.45$$

(1) เหตุการณ์ที่เหรียญออกหัว 1 ครั้ง หมายความว่า เหรียญจะออกหัวในการโยนครั้งที่ 1 ก็ได้ ครั้งที่ 2 ก็ได้ หรืออาจจะออกหัวในการโยนครั้งที่ 3 ก็ได้ ดังนั้น ความน่าจะเป็นของเหตุการณ์ดังกล่าวแทนได้ด้วยสัญลักษณ์ $P((H_1 \cap T_2 \cap T_3) \cup (T_1 \cap H_2 \cap T_3) \cup (T_1 \cap T_2 \cap H_3))$ แต่เนื่องจากเหตุการณ์ $H_1 \cap T_2 \cap T_3$ เหตุการณ์ $T_1 \cap H_2 \cap T_3$ และเหตุการณ์ $T_1 \cap T_2 \cap H_3$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน ดังนั้น

$$P((H_1 \cap T_2 \cap T_3) \cup (T_1 \cap H_2 \cap T_3) \cup (T_1 \cap T_2 \cap H_3)) = P(H_1 \cap T_2 \cap T_3) + P(T_1 \cap H_2 \cap T_3) + P(T_1 \cap T_2 \cap H_3)$$

และเนื่องจากการโยนออกหัวหรือก้อยในแต่ละครั้งของการโยนไม่มีผลต่อกัน ทำให้เหตุการณ์ H_i และ T_i เป็นอิสระต่อกัน ดังนั้น

$$P((H_1 \cap T_2 \cap T_3) \cup (T_1 \cap H_2 \cap T_3) \cup (T_1 \cap T_2 \cap H_3)) = P(H_1)P(T_2)P(T_3) + P(T_1)P(H_2)P(T_3) + P(T_1)P(T_2)P(H_3)$$

$$= (0.55)(0.45)(0.45) + (0.45)(0.55)(0.45) + (0.45)(0.45)(0.55)$$
$$= 0.3341$$

ดังนั้น การโยนเหรียญ 1 อัน 3 ครั้ง ความน่าจะเป็นที่เหรียญจะออกหัว 1 ครั้ง มีค่าเท่ากับ 0.3341

(2) เหตุการณ์ที่เหรียญออกหัว 2 ครั้ง หมายความว่า เหรียญจะออกหัวในการโยนครั้งที่ 1 และ 2 ก็ได้ ครั้งที่ 1 และ 3 ก็ได้ หรืออาจจะออกหัวในการโยนครั้งที่ 2 และ 3 ก็ได้ ดังนั้น ความน่าจะเป็นของเหตุการณ์ดังกล่าวแทนได้ด้วยสัญลักษณ์ $P((H_1 \cap H_2 \cap T_3) \cup (H_1 \cap T_2 \cap H_3) \cup (T_1 \cap H_2 \cap H_3))$ แต่เนื่องจากเหตุการณ์ $H_1 \cap H_2 \cap T_3$ เหตุการณ์ $H_1 \cap T_2 \cap H_3$ และเหตุการณ์ $T_1 \cap H_2 \cap H_3$ เป็นเหตุการณ์ที่ไม่เกิดร่วมกัน ดังนั้น

$$P((H_1 \cap H_2 \cap T_3) \cup (H_1 \cap T_2 \cap H_3) \cup (T_1 \cap H_2 \cap H_3)) = P(H_1 \cap H_2 \cap T_3) + P(H_1 \cap T_2 \cap H_3) + P(T_1 \cap H_2 \cap H_3)$$

และเนื่องจากการโยนออกหัวหรือก้อยในแต่ละครั้งของการโยนไม่มีผลต่อกัน ทำให้เหตุการณ์ H_i และ T_i เป็นอิสระต่อกัน ดังนั้น

$$P((H_1 \cap H_2 \cap T_3) \cup (H_1 \cap T_2 \cap H_3) \cup (T_1 \cap H_2 \cap H_3)) = P(H_1)P(H_2)P(T_3) + P(H_1)P(T_2)P(H_3) + P(T_1)P(H_2)P(H_3)$$
$$= (0.55)(0.55)(0.45) + (0.55)(0.45)(0.55) + (0.45)(0.55)(0.55)$$
$$= 0.4084$$

ดังนั้น การโยนเหรียญ 1 อัน 3 ครั้ง ความน่าจะเป็นที่เหรียญจะออกหัว 2 ครั้ง มีค่าเท่ากับ 0.4084

(3) ความน่าจะเป็นที่เหรียญจะออกหัวทั้ง 3 ครั้ง แทนได้ด้วย $P(H_1 \cap H_2 \cap H_3)$ เนื่องจากการโยนออกหัวในแต่ละครั้งของการโยนไม่มีผลต่อกัน ทำให้เหตุการณ์ H_i เป็นอิสระต่อกัน ดังนั้น

$$P(H_1 \cap H_2 \cap H_3) = P(H_1)P(H_2)P(H_3)$$
$$= (0.55)(0.55)(0.55)$$
$$= 0.1664$$

ดังนั้น การโยนเหรียญ 1 อัน 3 ครั้ง ความน่าจะเป็นที่เหรียญจะออกหัวทั้ง 3 ครั้ง มีค่าเท่ากับ 0.1664 ■

ตัวอย่างที่ 1.7.5 ถ้าสุ่มเลือกไพ่ 3 ใบจากสำรับซึ่งมีไพ่ทั้งหมด 52 ใบ โดยสุ่มทีละใบ จงหาความน่าจะเป็นที่จะสุ่มเลือกได้ไพ่โพธิ์แดงทั้ง 3 ใบ โดยที่

(1) เป็นการสุ่มแบบใส่คืน

(2) เป็นการสุ่มแบบไม่ใส่คืน

วิธีทำ กำหนดให้ A_i แทนเหตุการณ์ที่สุ่มเลือกได้ไพ่โพธิ์แดงในครั้งที่ i เมื่อ $i = 1, 2, 3$ ดังนั้น สิ่งที่ต้องการหาคือ $P(A_1 \cap A_2 \cap A_3)$

(1) เนื่องจากเป็นการสุ่มแบบใส่คืน ดังนั้น การสุ่มเลือกได้ไพ่โพธิ์แดงในแต่ละครั้งจึงเป็นอิสระต่อกัน (A_1, A_2 และ A_3 เป็นอิสระต่อกัน) พิจารณา

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$$
$$= \frac{13}{52} \cdot \frac{13}{52} \cdot \frac{13}{52} = 0.0156$$

ดังนั้น การสุ่มเลือกไพ่ 3 ใบจากสำรับซึ่งมีไพ่ทั้งหมด 52 ใบ โดยสุ่มทีละใบแบบใส่คืน ความน่าจะเป็นที่จะสุ่มเลือกได้ไพ่โพธิ์แดงทั้ง 3 ใบ มีค่าเท่ากับ 0.0156

(2) เนื่องจากเป็นการสุ่มแบบไม่ใส่คืน ดังนั้น การสุ่มเลือกได้ไพ่โพธิ์แดงในแต่ละครั้งจึงมีความเกี่ยวข้องกัน นั่นคือไม่เป็นอิสระต่อกัน (A_1, A_2 และ A_3 ไม่เป็นอิสระต่อกัน) พิจารณา

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)$$
$$= \frac{13}{52} \cdot \frac{12}{51} \cdot \frac{11}{50} = 0.0129$$

ดังนั้น การสุ่มเลือกไพ่ 3 ใบจากสำรับซึ่งมีไพ่ทั้งหมด 52 ใบ โดยสุ่มทีละใบแบบไม่ใส่คืน ความน่าจะเป็นที่จะสุ่มเลือกได้ไพ่โพธิ์แดงทั้ง 3 ใบ มีค่าเท่ากับ 0.0129 ■

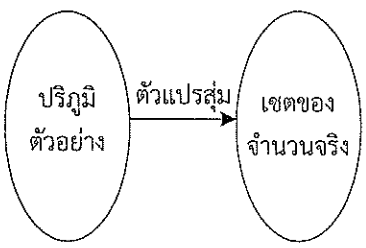
บทที่ 2
การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม (Probability Distribution of Random Variable)

ในบทที่ 2 นี้ได้กล่าวถึงตัวแปรสุ่ม และการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม โดยเริ่มต้นจากการอธิบายความหมายของตัวแปรสุ่ม จำแนกประเภทของตัวแปรสุ่ม และอธิบายถึงการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม 1 ตัว แล้วจึงขยายแนวความคิดไปเป็นการแจกแจงของตัวแปรสุ่ม 2 ตัว อีกทั้งยังกล่าวถึงความเป็นอิสระต่อกันของตัวแปรสุ่ม ค่าคาดหมาย ความแปรปรวน และส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่มด้วย

2.1 ความหมายของตัวแปรสุ่ม

<u>นิยามที่ 2.1.1</u> ***ตัวแปรสุ่ม (random variable)*** หมายถึง ฟังก์ชันที่ส่งจากปริภูมิตัวอย่าง S ไปยังเซตของจำนวนจริง $\mathbb{R}$

จากนิยามของตัวแปรสุ่ม โดเมนของตัวแปรสุ่ม คือ ปริภูมิตัวอย่าง ส่วนเรนจ์ของตัวแปรสุ่ม คือ สับเซตของจำนวนจริง ซึ่งสามารถอธิบายการส่งของตัวแปรสุ่มได้ตามรูปที่ 2.1.1



รูปที่ 2.1.1 การส่งของตัวแปรสุ่มจากปริภูมิตัวอย่างไปยังเซตของจำนวนจริง

ตัวแปรสุ่มนิยมเขียนแทนด้วยตัวอักษรภาษาอังกฤษตัวใหญ่ เช่น X, Y, Z และเขียนแทนค่าที่สอดคล้องกับตัวแปรสุ่มด้วยตัวอักษรภาษาอังกฤษตัวเล็ก เช่น x, y, z ดังนั้น $X(s) = x$ หมายความว่าค่า x (x เป็นจำนวนจริง) สัมพันธ์กับ s (s เป็นสมาชิกในปริภูมิตัวอย่าง) โดยมีตัวแปรสุ่ม X เป็นตัวกำหนดความสัมพันธ์

เมื่อกล่าวถึงตัวแปรสุ่ม จะกล่าวถึงสมาชิกที่อยู่ในเรนจ์ โดยที่ $\{X(s) : s \in S\}$ แทนเรนจ์ของตัวแปรสุ่ม ดังนั้นเพื่อความสะดวกจะเรียกสมาชิกทั้งหมดในเรนจ์ว่าค่าที่เป็นไปได้ทั้งหมดของตัวแปรสุ่ม X

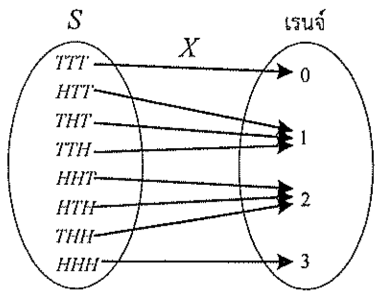
ตัวแปรสุ่มใช้เพื่ออธิบายถึงเหตุการณ์ต่าง ๆ โดยที่ $P(X = x)$ หมายถึงความน่าจะเป็นของเหตุการณ์ $\{s \in S : X(s) = x\}$

<u>ตัวอย่างที่ 2.1.1</u>

(1) พิจารณาการทดลองสุ่มโยนเหรียญสามเหรียญหนึ่งครั้ง เมื่อ H แทนเหตุการณ์ที่เหรียญออกหัว และ T แทนเหตุการณ์ที่เหรียญออกก้อย ดังนั้นปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

$$S = \{HHH,\ HHT,\ HTH,\ HTT,\ THH,\ THT,\ TTH,\ TTT\}$$

กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนเหรียญที่ออกหัวจากการโยนเหรียญสามเหรียญหนึ่งครั้ง ซึ่งการส่งของตัวแปรสุ่ม X เป็นไปตามรูปที่ 2.1.2



รูปที่ 2.1.2 การส่งของตัวแปรสุ่มที่แทนจำนวนเหรียญที่ออกหัวจากการโยนเหรียญสามเหรียญหนึ่งครั้ง

ดังนั้นฟังก์ชัน X มีนิยามดังนี้

$$X(HHH) = 3$$
$$X(HHT) = 2$$
$$X(HTH) = 2$$
$$X(HTT) = 1$$
$$X(THH) = 2$$
$$X(THT) = 1$$
$$X(TTH) = 1$$
$$X(TTT) = 0$$

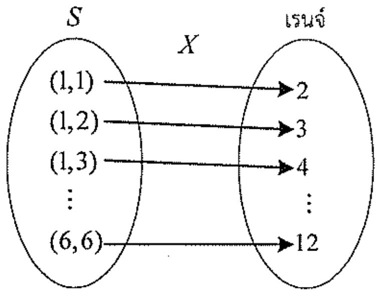
และเรนจ์ของตัวแปรสุ่ม X คือ

$$\{0,\ 1,\ 2,\ 3\}$$

(2) พิจารณาการทดลองสุ่มโยนลูกเต๋าสองลูกหนึ่งครั้ง เมื่อ (i, j) แทนเหตุการณ์ที่ลูกเต๋าลูกที่หนึ่งออกแต้ม i ส่วนลูกเต๋าลูกที่สองออกแต้ม j ดังนั้นปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

$$S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), \ldots, (6,6)\}$$

กำหนดให้ X เป็นตัวแปรสุ่มแทนผลรวมของแต้มบนลูกเต๋าทั้งสองจากการโยนลูกเต๋าสองลูกหนึ่งครั้ง ซึ่งการส่งของตัวแปรสุ่ม X เป็นไปตามรูปที่ 2.1.3



รูปที่ 2.1.3 การส่งของตัวแปรสุ่มที่แทนผลรวมของแต้มบนลูกเต๋าทั้งสองจากการโยนลูกเต๋าสองลูกหนึ่งครั้ง

ดังนั้นฟังก์ชัน X มีนิยามดังนี้

$$X\big((1,1)\big) = 2,\ X\big((1,2)\big) = 3,\ X\big((1,3)\big) = 4,\ \ldots,\ X\big((6,6)\big) = 12$$

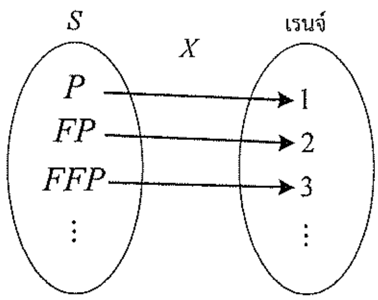
และเรนจ์ของตัวแปรสุ่ม X คือ

$$\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

(3) พิจารณาการทดลองสุ่มเลือกแบตเตอรีมาทดสอบจนกระทั่งได้แบตเตอรีที่ใช้งานได้ เมื่อ P แทนแบตเตอรีที่ใช้งานได้ และ F แทนแบตเตอรีที่ใช้งานไม่ได้ ดังนั้นปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

$$S = \{P, FP, FFP, \ldots\}$$

กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนครั้งที่เลือกแบตเตอรีจนกระทั่งได้แบตเตอรีที่ใช้งานได้ ซึ่งการส่งของตัวแปรสุ่ม X เป็นไปตามรูปที่ 2.1.4



รูปที่ 2.1.4 การส่งของตัวแปรสุ่มที่แทนจำนวนครั้งที่เลือกแบตเตอรีจนกระทั่งได้แบตเตอรีที่ใช้งานได้

ดังนั้นฟังก์ชัน X มีนิยามดังนี้

$$X(P)=1,\ X(FP)=2,\ X(FFP)=3,\ \ldots$$

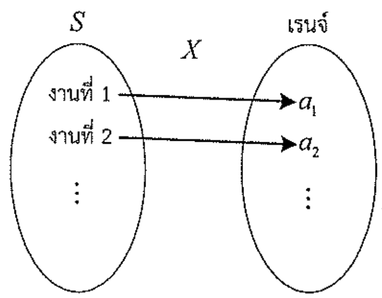
และเรนจ์ของตัวแปรสุ่ม X คือ

$$\{1,\ 2,\ 3,\ \ldots\}$$

(4) พิจารณาการทดลองสุ่มเลือกงานมาหนึ่งงานเพื่อให้มานพทำ จากนั้นจับเวลาที่มานพทำงานนี้แล้วเสร็จ ดังนั้นปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

$$S=\{\text{งานที่ } 1,\ \text{งานที่ } 2,\ \ldots\}$$

กำหนดให้ X เป็นตัวแปรสุ่มแทนระยะเวลาทั้งหมดที่ใช้ในการทำงานหนึ่งงานของมานพ (หน่วยเป็นนาที) ซึ่งการส่งของตัวแปรสุ่ม X เป็นไปตามรูปที่ 2.1.5



รูปที่ 2.1.5 การส่งของตัวแปรสุ่มที่แทนระยะเวลาทั้งหมดที่ใช้ในการทำงานหนึ่งงานของมานพ

ดังนั้นฟังก์ชัน X มีนิยามดังนี้

$$X(\text{งานที่ } 1)=a_1,\ X(\text{งานที่ } 2)=a_2,\ \ldots$$

เมื่อ a_i แทนระยะเวลาทั้งหมดที่มานพใช้ในการทำงานที่ i และเรนจ์ของตัวแปรสุ่ม X คือ

$$\{x : x \geq 0\}=[0,\infty)$$

■

2.2 ประเภทของตัวแปรสุ่ม

เมื่อพิจารณาเรนจ์ของตัวแปรสุ่มจากตัวอย่างที่ 2.1.1 จะพบความแตกต่างซึ่งทำให้สามารถจำแนกตัวแปรสุ่มออกได้เป็น 2 ประเภท คือ ตัวแปรสุ่มแบบไม่ต่อเนื่อง และตัวแปรสุ่มแบบต่อเนื่อง ตามนิยามต่อไปนี้

<u>นิยามที่ 2.2.1</u> ***ตัวแปรสุ่มแบบไม่ต่อเนื่อง (discrete random variable)*** หมายถึง ตัวแปรสุ่มที่มีเรนจ์เป็นเซตจำกัดหรือเซตอนันต์ที่นับได้

ถ้า X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ดังนั้น ค่าที่เป็นไปได้ทั้งหมดของ X คือ $x_1, x_2, \ldots, x_n$ (หรือ $x_1, x_2, \ldots, x_n, \ldots$) โดยนิยมเขียนเรียงลำดับดังนี้ $x_1 < x_2 < \ldots < x_n$ (หรือ $x_1 < x_2 < \ldots < x_n < \ldots$) และค่าที่เป็นไปได้ของ X ไม่จำเป็นต้องเป็นจำนวนนับเสมอไป อาจจะเป็นทศนิยมหรือเศษส่วนก็ได้

<u>**นิยามที่ 2.2.2**</u> ***ตัวแปรสุ่มแบบต่อเนื่อง (continuous random variable)*** หมายถึง ตัวแปรสุ่มที่มีเรนจ์เป็นช่วงที่เป็นสับเซตของจำนวนจริง

ถ้า X เป็นตัวแปรสุ่มแบบต่อเนื่อง ดังนั้น ค่าที่เป็นไปได้ทั้งหมดของ X จะเป็นสมาชิกของช่วงใดช่วงหนึ่งต่อไปนี้ $(a,b),[a,b],[a,b),(a,b],[a,\infty),(-\infty,b],(a,\infty),(-\infty,b)$ หรือ $(-\infty,\infty)$ เมื่อ a และ b เป็นจำนวนจริงใด ๆ

<u>ตัวอย่างที่ 2.2.1</u>

(1) กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนเหรียญที่ออกหัวจากการโยนเหรียญสามเหรียญหนึ่งครั้ง ดังนั้นเรนจ์ของตัวแปรสุ่ม X คือ

$$\{0,\ 1,\ 2,\ 3\}$$

เนื่องจากเรนจ์ของตัวแปรสุ่ม X มีสมาชิกทั้งหมด 4 ค่า นั่นคือเรนจ์ของตัวแปรสุ่ม X เป็นเซตจำกัด ดังนั้น X จึงเป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง

(2) กำหนดให้ X เป็นตัวแปรสุ่มแทนระยะทางที่มานะเดินในแต่ละวัน (หน่วยเป็นเมตร) ดังนั้นเรนจ์ของตัวแปรสุ่ม X คือ

$$\{x : x \geq 0\} = [0,\infty)$$

เนื่องจากเรนจ์ของตัวแปรสุ่ม X เป็นช่วงที่เป็นสับเซตของจำนวนจริง ดังนั้น X จึงเป็นตัวแปรสุ่มแบบต่อเนื่อง

(3) กำหนดให้ X เป็นตัวแปรสุ่มแทนน้ำหนักของนักศึกษาชั้นปีที่ 1 ของมหาวิทยาลัยเทคโนโลยีสุรนารีทั้งหมด (หน่วยเป็นกิโลกรัม) ดังนั้นเรนจ์ของตัวแปรสุ่ม X คือ

$$\{x : x \geq 0\} = [0,\infty)$$

เนื่องจากเรนจ์ของตัวแปรสุ่ม X เป็นช่วงที่เป็นสับเซตของจำนวนจริง ดังนั้น X จึงเป็นตัวแปรสุ่มแบบต่อเนื่อง

(4) กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนต้นลำไยในพื้นที่ 1 ไร่ ดังนั้นเรนจ์ของตัวแปรสุ่ม X คือ

$$\{0,\ 1,\ 2,\ \ldots\}$$

เนื่องจากเรนจ์ของตัวแปรสุ่ม X เป็นเซตอนันต์ที่นับได้ ดังนั้น X จึงเป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ■

2.3 การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม 1 ตัว

ในหัวข้อนี้จะกล่าวถึงการแจกแจงความน่าจะเป็นของตัวแปรสุ่มเพียงหนึ่งตัวเท่านั้น ซึ่งประกอบไปด้วยฟังก์ชันต่าง ๆ ที่สำคัญของตัวแปรสุ่ม สมบัติที่สำคัญของฟังก์ชันเหล่านี้ รวมถึงอธิบายความสัมพันธ์ของฟังก์ชันต่าง ๆ ดังนั้นเพื่อความสะดวกในหัวข้อนี้จะกล่าวถึงตัวแปรสุ่ม X

2.3.1 ฟังก์ชันความน่าจะเป็น

ในหัวข้อนี้จะกล่าวถึงฟังก์ชันที่สำคัญฟังก์ชันหนึ่งของตัวแปรสุ่ม นั่นคือ ***ฟังก์ชันความน่าจะเป็น (probability function หรือ pf)*** ซึ่งเขียนแทนด้วยสัญลักษณ์ $f_X(x)$ หมายถึง ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ที่ค่า $X = x$ แต่เนื่องจากกล่าวถึงเฉพาะตัวแปรสุ่ม X เพียงตัวเดียว ดังนั้นจึงละตัวแปร X และเขียนเพียง $f(x)$ เท่านั้น

> นิยามที่ 2.3.1 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม หมายถึง ฟังก์ชันที่ส่งจากเซตของจำนวนจริงไปยังเซตของจำนวนจริง

สำหรับฟังก์ชันความน่าจะเป็นนี้จะนิยามแตกต่างกันไปในตัวแปรสุ่มแบบไม่ต่อเนื่อง และตัวแปรสุ่มแบบต่อเนื่อง ดังนี้

2.3.1.1 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง

> นิยามที่ 2.3.2 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่องที่ค่า ๆ หนึ่ง มีค่าเท่ากับความน่าจะเป็นของเหตุการณ์ที่ตัวแปรสุ่มมีค่าเท่ากับค่า ๆ นั้น นั่นคือ ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง X ที่ค่า x นิยามดังนี้
>
> $$f(x) = P(X = x)$$
>
> สำหรับทุกค่า $x \in \mathbb{R}$ (หรือสำหรับทุกค่า x ที่เป็นจำนวนจริง)
>
> ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง เรียกได้อีกอย่างหนึ่งว่า ***ฟังก์ชันความน่าจะเป็นของมวล (probability mass function หรือ pmf)***

สมบัติที่สำคัญของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง

1. ฟังก์ชันความน่าจะเป็นมีค่าตั้งแต่ศูนย์ถึงหนึ่งเสมอ สำหรับทุกจำนวนจริง x นั่นคือ

$$0 \le f(x) \le 1$$

สำหรับทุกค่า $x \in \mathbb{R}$

2. ผลรวมของฟังก์ชันความน่าจะเป็นสำหรับทุกจำนวนจริง x มีค่าเท่ากับหนึ่งเสมอ นั่นคือ

$$\sum_{x \in \mathbb{R}} f(x) = 1$$

หรืออาจกล่าวได้ว่า ผลรวมของฟังก์ชันความน่าจะเป็นสำหรับทุกค่าที่เป็นไปได้ของตัวแปรสุ่ม X มีค่าเท่ากับหนึ่งเสมอ นั่นคือ

$$\sum_{x} f(x) = 1$$

สำหรับทุกค่า x ที่เป็นสมาชิกของเรนจ์ของตัวแปรสุ่ม X

3. ความน่าจะเป็นของเหตุการณ์ A (เหตุการณ์ A เป็นเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม X) มีค่าเท่ากับผลรวมของฟังก์ชันความน่าจะเป็นสำหรับทุกค่าของตัวแปรสุ่ม X ที่เป็นสมาชิกของเหตุการณ์ A นั่นคือ

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

4. ฟังก์ชันความน่าจะเป็นสำหรับค่า x ที่ไม่เป็นสมาชิกของเรนจ์ของตัวแปรสุ่ม X มีค่าเท่ากับศูนย์เสมอ นั่นคือ สำหรับทุกค่า x ที่ไม่เป็นสมาชิกของเรนจ์ของตัวแปรสุ่ม X

$$f(x) = 0$$

<u>ตัวอย่างที่ 2.3.1</u> สายการบินแห่งหนึ่งต้องการทราบสถิติการร้องเรียนของลูกค้า จึงทำการเก็บข้อมูลจำนวนครั้งที่ลูกค้าร้องเรียนในแต่ละวันเป็นจำนวน 30 วัน ซึ่งผลเป็นไปตามตารางต่อไปนี้

จำนวนครั้งที่ลูกค้าร้องเรียนต่อวัน	0	1	2	3	4	5
จำนวนวัน	10	8	5	4	2	1

จงหา

(1) ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่แทนจำนวนครั้งที่ลูกค้าร้องเรียนในแต่ละวันของสายการบินแห่งนี้

(2) ความน่าจะเป็นที่สายการบินแห่งนี้จะถูกลูกค้าร้องเรียนมากกว่า 3 ครั้งต่อวัน

(3) ความน่าจะเป็นที่สายการบินแห่งนี้จะถูกลูกค้าร้องเรียนตั้งแต่ 3 ครั้งขึ้นไปต่อวัน

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนครั้งที่ลูกค้าร้องเรียนในแต่ละวันของสายการบินแห่งนี้ ดังนั้นค่าที่เป็นไปได้ทั้งหมดของ X คือ 0, 1, 2, 3, 4, 5 เพราะฉะนั้น X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง และสามารถหาค่าความน่าจะเป็นของแต่ละค่าของ X ได้ดังตารางต่อไปนี้

จำนวนครั้งที่ลูกค้าร้องเรียนต่อวัน	0	1	2	3	4	5
จำนวนวัน	10	8	5	4	2	1
ความน่าจะเป็น	10/30	8/30	5/30	4/30	2/30	1/30

(1) จากนิยามฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่แทนจำนวนครั้งที่ลูกค้าร้องเรียนในแต่ละวันของสายการบินแห่งนี้คือ

$$f(x) = P(X = x)$$

สำหรับทุกค่า $x \in \mathbb{R}$ ดังนั้นหาค่า $f(x)$ เมื่อ $x = 0,1,2,3,4,5$ ได้ดังนี้

$$f(0) = P(X = 0) = \frac{10}{30}$$

$$f(1) = P(X = 1) = \frac{8}{30}$$

$$f(2) = P(X = 2) = \frac{5}{30}$$

$$f(3) = P(X = 3) = \frac{4}{30}$$

$$f(4) = P(X = 4) = \frac{2}{30}$$

$$f(5) = P(X = 5) = \frac{1}{30}$$

ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ที่แทนจำนวนครั้งที่ลูกค้าร้องเรียนในแต่ละวันของสายการบินแห่งนี้เป็นไปตามสมการ

$$f(x) = \begin{cases} \frac{10}{30}, & x = 0 \\ \frac{8}{30}, & x = 1 \\ \frac{5}{30}, & x = 2 \\ \frac{4}{30}, & x = 3 \\ \frac{2}{30}, & x = 4 \\ \frac{1}{30}, & x = 5 \\ 0, & x \in \mathbb{R} - \{0,1,2,3,4,5\} \end{cases}$$

(2) ความน่าจะเป็นที่สายการบินแห่งนี้จะถูกลูกค้าร้องเรียนมากกว่า 3 ครั้งต่อวัน คือ $P(X > 3)$ จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง

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

พิจารณา $A = \{X > 3\} = \{4,5\}$ ดังนั้น

$$P(X > 3) = \sum_{X \in \{4,5\}} f(x) = f(4) + f(5) = \frac{2}{30} + \frac{1}{30} = \frac{3}{30} = \frac{1}{10} = 0.1$$

ดังนั้นความน่าจะเป็นที่สายการบินแห่งนี้จะถูกลูกค้าร้องเรียนมากกว่า 3 ครั้งต่อวันมีค่าเท่ากับ 0.1

(3) ความน่าจะเป็นที่สายการบินแห่งนี้จะถูกลูกค้าร้องเรียนตั้งแต่ 3 ครั้งขึ้นไปต่อวัน คือ $P(X \geq 3)$ ใช้สมบัติเดียวกันกับข้อ (2) โดยพิจารณา $A = \{X \geq 3\} = \{3,4,5\}$ ดังนั้น

$$P(X \geq 3) = \sum_{X \in \{3,4,5\}} f(x) = f(3) + f(4) + f(5) = \frac{4}{30} + \frac{2}{30} + \frac{1}{30} = \frac{7}{30}$$

ดังนั้นความน่าจะเป็นที่สายการบินแห่งนี้จะถูกลูกค้าร้องเรียนตั้งแต่ 3 ครั้งขึ้นไปต่อวันมีค่าเท่ากับ $\frac{7}{30}$ ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 2.3.1 ข้อ (2) และ (3) จะเห็นว่า $P(X > 3) \neq P(X \geq 3)$

<u>ตัวอย่างที่ 2.3.2</u> การสำรวจความยาวของชิ้นส่วนต่าง ๆ ที่ผลิตจากโรงงานแห่งหนึ่งจำนวน 115 ชิ้น สามารถจำแนกชิ้นส่วนต่าง ๆ ตามความยาวได้ตามตารางนี้

ความยาว (มิลลิเมตร)	4.9	5.0	5.1	5.2	5.3	5.4	5.5	5.6
จำนวนชิ้นส่วน	1	3	10	25	40	18	16	2

จงหา

(1) ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่แทนความยาวของชิ้นส่วนที่ผลิตจากโรงงานแห่งนี้

(2) ความน่าจะเป็นที่สุ่มได้ชิ้นส่วนที่มีความยาวไม่เกิน 5.1 มิลลิเมตร

(3) ความน่าจะเป็นที่สุ่มได้ชิ้นส่วนที่มีความยาวระหว่าง 4.95 - 5.35 มิลลิเมตร

วิธีทำ กำหนดให้ X แทนความยาวของชิ้นส่วนที่ผลิตจากโรงงานแห่งนี้ จากตารางที่แสดงข้างต้น ค่าที่เป็นไปได้ทั้งหมดของ X คือ 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6 ดังนั้น X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง และสามารถหาค่าความน่าจะเป็นของแต่ละค่าของ X ได้ดังตารางต่อไปนี้

ความยาว (มิลลิเมตร)	4.9	5.0	5.1	5.2	5.3	5.4	5.5	5.6
จำนวนชิ้นส่วน	1	3	10	25	40	18	16	2
ความน่าจะเป็น	1/115	3/115	10/115	25/115	40/115	18/115	16/115	2/115

(1) จากนิยามของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่แทนความยาวของชิ้นส่วนที่ผลิตจากโรงงานแห่งนี้คือ

$$f(x) = P(X = x)$$

สำหรับทุกค่า $x \in \mathbb{R}$ ดังนั้นหาค่า $f(x)$ เมื่อ $x = 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6$ ได้ดังนี้

$$f(4.9) = P(X = 4.9) = \frac{1}{115}$$

$$f(5.0) = P(X = 5.0) = \frac{3}{115}$$

$$f(5.1) = P(X = 5.1) = \frac{10}{115}$$
$$f(5.2) = P(X = 5.2) = \frac{25}{115}$$
$$f(5.3) = P(X = 5.3) = \frac{40}{115}$$
$$f(5.4) = P(X = 5.4) = \frac{18}{115}$$
$$f(5.5) = P(X = 5.5) = \frac{16}{115}$$
$$f(5.6) = P(X = 5.6) = \frac{2}{115}$$

ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ที่แทนความยาวของชิ้นส่วนที่ผลิตจากโรงงานแห่งนี้เป็นไปตามสมการ

$$f(x) = \begin{cases} \dfrac{1}{115}, & x = 4.9 \\ \dfrac{3}{115}, & x = 5.0 \\ \dfrac{10}{115}, & x = 5.1 \\ \dfrac{25}{115}, & x = 5.2 \\ \dfrac{40}{115}, & x = 5.3 \\ \dfrac{18}{115}, & x = 5.4 \\ \dfrac{16}{115}, & x = 5.5 \\ \dfrac{2}{115}, & x = 5.6 \\ 0, & x \in \mathbb{R} - \{4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6\} \end{cases}$$

(2) ความน่าจะเป็นที่สุ่มได้ชิ้นส่วนที่มีความยาวไม่เกิน 5.1 มิลลิเมตร คือ $P(X \le 5.1)$ จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง

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

พิจารณา $A = \{X \le 5.1\} = \{4.9, 5.0, 5.1\}$ ดังนั้น

$$P(X \le 5.1) = \sum_{X \in \{4.9, 5.0, 5.1\}} f(x) = f(4.9) + f(5.0) + f(5.1) = \frac{1}{115} + \frac{3}{115} + \frac{10}{115} = \frac{14}{115}$$

ดังนั้นความน่าจะเป็นที่สุ่มได้ชิ้นส่วนที่มีความยาวไม่เกิน 5.1 มิลลิเมตร มีค่าเท่ากับ $\frac{14}{115}$

(3) ความน่าจะเป็นที่สุ่มได้ชิ้นส่วนที่มีความยาวระหว่าง 4.95 - 5.35 มิลลิเมตร มีค่าเท่ากับ $P(4.95 < X < 5.35)$ ใช้สมบัติเดียวกันกับข้อ (2) โดยพิจารณา $A = \{4.95 < X < 5.35\} = \{5.0, 5.1, 5.2, 5.3\}$ ดังนั้น

$$P(4.95 < X < 5.35) = \sum_{X \in \{5.0, 5.1, 5.2, 5.3\}} f(x) = f(5.0) + f(5.1) + f(5.2) + f(5.3)$$

$$= \frac{3}{115} + \frac{10}{115} + \frac{25}{115} + \frac{40}{115} = \frac{78}{115}$$

ดังนั้นความน่าจะเป็นที่สุ่มได้ชิ้นส่วนที่มีความยาวระหว่าง 4.95 – 5.35 มิลลิเมตร มีค่าเท่ากับ $\frac{78}{115}$ ■

2.3.1.2 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง

<u>นิยามที่ 2.3.3</u> กำหนดให้ X เป็นตัวแปรสุ่มแบบต่อเนื่อง พิจารณาฟังก์ชัน f ที่ส่งจากเซตของจำนวนจริงไปยังเซตของจำนวนจริง โดยที่

$$f(x) \ge 0$$

สำหรับทุกค่า $x \in \mathbb{R}$ และ

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

เรียกฟังก์ชัน f ซึ่งสอดคล้องกับเงื่อนไขเหล่านี้ว่า ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง เรียกได้อีกอย่างหนึ่งว่า ***ฟังก์ชันความน่าจะเป็นของความหนาแน่น (probability density function หรือ pdf)***

สมบัติที่สำคัญของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง

1. ฟังก์ชันความน่าจะเป็นมีค่ามากกว่าหรือเท่ากับศูนย์เสมอ สำหรับทุกจำนวนจริง x นั่นคือ

$$f(x) \ge 0$$

สำหรับทุกค่า $x \in \mathbb{R}$

2. อินทิเกรตฟังก์ชันความน่าจะเป็นในช่วง $(-\infty, \infty)$ มีค่าเท่ากับหนึ่งเสมอ นั่นคือ

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

3. ความน่าจะเป็นของเหตุการณ์ $A = \{a \le X \le b\}$ มีค่าเท่ากับอินทิเกรตฟังก์ชันความน่าจะเป็นในช่วง (a, b) นั่นคือ

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

โดยที่ $-\infty < a \le b < \infty$

4. เหตุการณ์ที่ตัวแปรสุ่มมีค่าเท่ากับค่า ๆ หนึ่งจะมีความน่าจะเป็นเท่ากับศูนย์เสมอ นั่นคือ

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

สำหรับทุกค่า $a \in \mathbb{R}$

5. ความน่าจะเป็นของเหตุการณ์ที่ตัวแปรสุ่มมีค่าดังต่อไปนี้มีค่าเท่ากัน นั่นคือ

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

สำหรับทุกค่า $a, b \in \mathbb{R}$ และ $a \le b$

6. เหตุการณ์ที่ตัวแปรสุ่มมีค่าน้อยกว่าหรือเท่ากับ (หรือน้อยกว่า) ค่า ๆ หนึ่ง สามารถหาค่าความน่าจะเป็นได้ดังนี้

$$P(X \le a) = P(X < a) = \int_{-\infty}^a f(x)dx$$

สำหรับทุกค่า $a \in \mathbb{R}$

7. เหตุการณ์ที่ตัวแปรสุ่มมีค่ามากกว่าหรือเท่ากับ (หรือมากกว่า) ค่า ๆ หนึ่ง สามารถหาค่าความน่าจะเป็นได้ดังนี้

$$P(X \ge a) = P(X > a) = \int_a^\infty f(x)dx$$

สำหรับทุกค่า $a \in \mathbb{R}$

จากสมบัติข้อที่ 4 - 7 จะเห็นว่าเป็นจริงสำหรับกรณีที่ตัวแปรสุ่มเป็นแบบต่อเนื่องเท่านั้น ส่วนในกรณีที่ตัวแปรสุ่มเป็นแบบไม่ต่อเนื่องสมบัติทั้งสี่ข้อนี้ไม่เป็นจริง

<u>ตัวอย่างที่ 2.3.3</u> กำหนดให้ตัวแปรสุ่มแบบต่อเนื่อง X มีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x) = \begin{cases} \frac{1}{2} - ax, & x \in (0,4) \\ 0, & x \notin (0,4) \end{cases}$$

เมื่อ a เป็นค่าคงที่ จงหา

(1) ค่า a

(2) $P(1 < X < 3)$

(3) $P(1 < X < 4)$

(4) $P(1 < X < 5)$

วิธีทำ (1) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง

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

ดังนั้น

$$\int_{-\infty}^{0} 0dx + \int_{0}^{4}\left(\frac{1}{2} - ax\right)dx + \int_{4}^{\infty} 0dx = 1$$

$$\int_{0}^{4}\left(\frac{1}{2} - ax\right)dx = 1$$

$$\left[\frac{1}{2}x - \frac{ax^2}{2}\right]_{x=0}^{x=4} = 1$$

$$2 - 8a = 1$$

$$a = \frac{1}{8}$$

แทนค่า $a = \frac{1}{8}$ ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง X คือ

$$f(x) = \begin{cases} \frac{1}{2} - \frac{1}{8}x, & x \in (0,4) \\ 0, & x \notin (0,4) \end{cases}$$

(2) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$\begin{aligned} P(1 < X < 3) &= \int_{1}^{3} f(x)dx \\ &= \int_{1}^{3}\left(\frac{1}{2} - \frac{1}{8}x\right)dx \\ &= \left[\frac{x}{2} - \frac{x^2}{16}\right]_{x=1}^{x=3} \\ &= \left(\frac{3}{2} - \frac{9}{16}\right) - \left(\frac{1}{2} - \frac{1}{16}\right) \\ &= \frac{15}{16} - \frac{7}{16} = \frac{8}{16} = \frac{1}{2} = 0.5 \end{aligned}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์ที่ $1 < X < 3$ มีค่าเท่ากับ 0.5

(3) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$P(1 < X < 4) = \int_{1}^{4} f(x)dx$$

$$= \int_1^4 \left(\frac{1}{2} - \frac{1}{8}x\right) dx$$

$$= \left[\frac{x}{2} - \frac{x^2}{16}\right]_{x=1}^{x=4}$$

$$= \left(\frac{4}{2} - \frac{16}{16}\right) - \left(\frac{1}{2} - \frac{1}{16}\right)$$

$$= 1 - \frac{7}{16} = \frac{9}{16}$$

ดังนั้นความน่าจะของเหตุการณ์ที่ $1 < X < 4$ มีค่าเท่ากับ $\frac{9}{16}$

(4) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$P(1 < X < 5) = \int_1^5 f(x)dx$$

$$= \int_1^4 \left(\frac{1}{2} - \frac{1}{8}x\right) dx + \int_4^5 0dx$$

$$= \int_1^4 \left(\frac{1}{2} - \frac{1}{8}x\right) dx$$

$$= \left[\frac{x}{2} - \frac{x^2}{16}\right]_{x=1}^{x=4}$$

$$= \left(\frac{4}{2} - \frac{16}{16}\right) - \left(\frac{1}{2} - \frac{1}{16}\right)$$

$$= 1 - \frac{7}{16} = \frac{9}{16}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์ที่ $1 < X < 5$ มีค่าเท่ากับ $\frac{9}{16}$ ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 2.3.3 ถ้าทราบคำตอบข้อ (3) เราจะสามารถหาคำตอบของข้อ (4) ได้ทันที โดยไม่จำเป็นต้องคำนวณซ้ำอีก

<u>ตัวอย่างที่ 2.3.4</u> กำหนดให้ตัวแปรสุ่มแบบต่อเนื่อง X มีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x)=\begin{cases} \dfrac{x}{6}, & 0\le x\le 2 \\ A(x-2), & 2<x\le 4 \\ 0, & x\notin[0,4] \end{cases}$$

เมื่อ A เป็นค่าคงที่ จงหา

(1) ค่า A

(2) $P(1<X<3)$

(3) $P(X\ge 3)$

(4) $P(X\le 1)$

วิธีทำ (1) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง

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

ดังนั้น

$$\int_{-\infty}^{0} 0dx+\int_{0}^{2}\frac{x}{6}dx+\int_{2}^{4}A(x-2)dx+\int_{4}^{\infty}0dx=1$$

$$\int_{0}^{2}\frac{x}{6}dx+\int_{2}^{4}A(x-2)dx=1$$

$$\left[\frac{x^2}{12}\right]_{x=0}^{x=2}+\left[A\left(\frac{x^2}{2}-2x\right)\right]_{x=2}^{x=4}=1$$

$$\frac{1}{3}+A\big((8-8)-(2-4)\big)=1$$

$$\frac{1}{3}+2A=1$$

$$A=\frac{1}{3}$$

แทนค่า $A=\dfrac{1}{3}$ ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง X คือ

$$f(x)=\begin{cases} \dfrac{x}{6}, & 0\le x\le 2 \\ \dfrac{1}{3}(x-2), & 2<x\le 4 \\ 0, & x\notin[0,4] \end{cases}$$

(2) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$\begin{aligned}P(1<X<3) &= \int_1^3 f(x)dx \\ &= \int_1^2 \frac{x}{6}dx + \int_2^3 \frac{1}{3}(x-2)dx \\ &= \left[\frac{x^2}{12}\right]_{x=1}^{x=2} + \left[\frac{1}{3}\left(\frac{x^2}{2}-2x\right)\right]_{x=2}^{x=3} \\ &= \left(\frac{1}{3}-\frac{1}{12}\right)+\frac{1}{3}\left\{\left(\frac{9}{2}-6\right)-(2-4)\right\} \\ &= \frac{3}{12}+\frac{1}{3}\left(\frac{1}{2}\right)=\frac{5}{12}\end{aligned}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์ที่ $1<X<3$ มีค่าเท่ากับ $\frac{5}{12}$

(3) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$\begin{aligned}P(X\geq 3) &= \int_3^\infty f(x)dx = \int_3^4 \frac{1}{3}(x-2)dx = \frac{1}{3}\left[\frac{x^2}{2}-2x\right]_{x=3}^{x=4} \\ &= \frac{1}{3}\left\{(8-8)-\left(\frac{9}{2}-6\right)\right\}=\frac{1}{3}\left(\frac{3}{2}\right)=\frac{1}{2}=0.5\end{aligned}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์ที่ $X\geq 3$ มีค่าเท่ากับ 0.5

(4) จากสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$P(X\leq 1)=\int_{-\infty}^1 f(x)dx = \int_0^1 \frac{x}{6}dx = \left[\frac{x^2}{12}\right]_{x=0}^{x=1} = \frac{1}{12}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์ที่ $X\leq 1$ มีค่าเท่ากับ $\frac{1}{12}$ ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 2.3.4 ข้อ (2) – (3) จำเป็นจะต้องทราบค่า A ก่อนถึงจะสามารถหาคำตอบได้ ส่วนคำถามข้อ (4) เราสามารถหาคำตอบได้เลยโดยที่ไม่จำเป็นต้องทราบค่า A

2.3.2 ฟังก์ชันการแจกแจงสะสม

ในหัวข้อนี้จะกล่าวถึงฟังก์ชันที่สำคัญอีกฟังก์ชันหนึ่งของตัวแปรสุ่ม นั่นคือ ***ฟังก์ชันการแจกแจงสะสม (cumulative distribution function หรือ cdf)*** ซึ่งเขียนแทนด้วยสัญลักษณ์ $F_X(x)$ หมายถึง ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X ที่ค่า $X = x$ เพื่อความสะดวกจะละ X และเขียน $F(x)$ แทน $F_X(x)$ สำหรับฟังก์ชันการแจกแจงสะสมมีนิยาม ดังนี้

<u>นิยามที่ 2.3.4.</u> *ฟังก์ชันการแจกแจงสะสม (cumulative distribution function)* ของตัวแปรสุ่ม X หมายถึง ฟังก์ชันที่ส่งจากเซตของจำนวนจริงไปยัง $[0,1]$ โดยบอกถึงความน่าจะเป็นสะสม เมื่อตัวแปรสุ่ม X มีค่าน้อยกว่าค่า ๆ หนึ่ง นั่นคือ

$$F(x) = P(X \le x)$$

สมบัติที่สำคัญของฟังก์ชันการแจกแจงสะสม

1. ฟังก์ชันการแจกแจงสะสมมีค่าตั้งแต่ศูนย์ถึงหนึ่งเสมอ สำหรับทุกจำนวนจริง x เนื่องจากความน่าจะเป็นมีค่าตั้งแต่ศูนย์ถึงหนึ่ง นั่นคือ

$$0 \le F(x) \le 1$$

สำหรับทุกค่า $x \in \mathbb{R}$

2. ฟังก์ชันการแจกแจงสะสมเป็นฟังก์ชันไม่ลด (non-decreasing function) หมายความว่า เมื่อ x มีค่ามากขึ้น ฟังก์ชันการแจกแจงสะสมจะมีค่ามากขึ้นหรืออาจมีค่าเท่าเดิม แต่จะไม่มีค่าลดลง นั่นคือ

$$F(x_1) \le F(x_2)$$

สำหรับทุกค่า $x_1, x_2 \in \mathbb{R}$ และ $x_1 \le x_2$

3. ฟังก์ชันการแจกแจงสะสมเป็นฟังก์ชันต่อเนื่องจากทางขวา นั่นคือ

$$\lim_{x \to a^+} F(x) = F(a)$$

4. ฟังก์ชันการแจกแจงสะสมมีค่าลู่เข้าสู่ศูนย์เมื่อ x มีค่าน้อย ๆ และมีค่าลู่เข้าสู่หนึ่งเมื่อ x มีค่ามาก ๆ นั่นคือ

$$\lim_{x \to -\infty} F(x) = 0 \text{ และ } \lim_{x \to \infty} F(x) = 1$$

2.3.2.1 ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มแบบไม่ต่อเนื่อง

<u>นิยามที่ 2.3.5</u> ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มแบบไม่ต่อเนื่องที่ค่า ๆ หนึ่ง มีค่าเท่ากับผลรวมของฟังก์ชันความน่าจะเป็นของจำนวนจริงทุกจำนวนที่น้อยกว่าหรือเท่ากับค่า ๆ นั้น นั่นคือ

$$F(x) = P(X \le x) = \sum_{a \le x} f(a)$$

<u>**ตัวอย่างที่ 2.3.5**</u> กำหนดให้ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง โดยมีค่าที่เป็นไปได้ทั้งหมดคือ 0, 1, 2 และ 4 จงหาฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X เมื่อ x มีค่าเท่ากับ 0, 1, 2, 3 และ 4 ในเทอมของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X

วิธีทำ จากนิยามฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มแบบไม่ต่อเนื่องจะได้ว่า

$$F(x)=\sum_{a\leq x} f(a)$$

ดังนั้นเราสามารถหาฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X ที่ค่า 0, 1, 2, 3 และ 4 ในเทอมของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ได้ดังนี้

$$F(0)=\sum_{a\leq 0} f(a)=f(0)$$

$$F(1)=\sum_{a\leq 1} f(a)=f(0)+f(1)$$

$$F(2)=\sum_{a\leq 2} f(a)=f(0)+f(1)+f(2)$$

$$F(3)=\sum_{a\leq 3} f(a)=f(0)+f(1)+f(2)$$

$$F(4)=\sum_{a\leq 4} f(a)=f(0)+f(1)+f(2)+f(4)$$

■

<u>**ตัวอย่างที่ 2.3.6**</u> ถ้าตัวแปรสุ่มแบบไม่ต่อเนื่อง X มีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x)=\begin{cases}\dfrac{x}{6}, & x=1,2,3\\ 0, & x\neq 1,2,3\end{cases}$$

จงหา

(1) $F(2)$

(2) $F(1.5)$

(3) ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X

วิธีทำ จากนิยามฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มแบบไม่ต่อเนื่องจะได้ว่า

$$F(x)=\sum_{a\leq x} f(a)$$

(1) $F(2)=\sum_{a\leq 2} f(a)=f(1)+f(2)=\frac{1}{6}+\frac{2}{6}=\frac{3}{6}=\frac{1}{2}$

(2) $F(1.5)=\sum_{a\leq 1.5} f(a)=f(1)=\frac{1}{6}$

(3) เนื่องจากฟังก์ชันการแจกแจงสะสมเป็นฟังก์ชันไม่ลด ดังนั้นเราจะแยกพิจารณาเป็นช่วง ๆ ดังนี้

เมื่อ $x<1$:

$$F(x)=\sum_{a\leq x} f(a)=0$$

เมื่อ $1 \le x < 2$:

$$F(x) = \sum_{a \le x} f(a) = f(1) = \frac{1}{6}$$

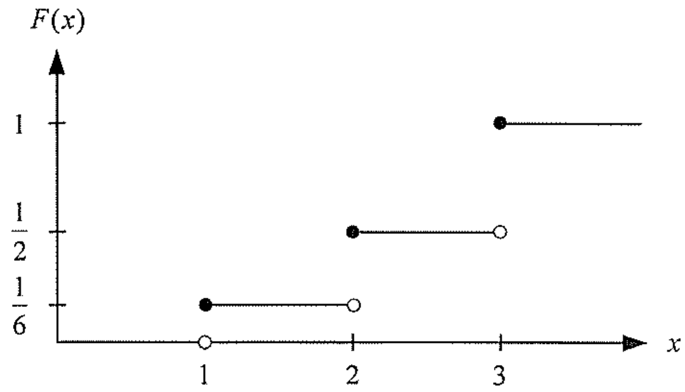
เมื่อ $2 \le x < 3$:

$$F(x) = \sum_{a \le x} f(a) = f(1) + f(2) = \frac{1}{6} + \frac{2}{6} = \frac{3}{6} = \frac{1}{2}$$

เมื่อ $x \ge 3$:

$$F(x) = \sum_{a \le x} f(a) = f(1) + f(2) + f(3) = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} = \frac{6}{6} = 1$$

กราฟของฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X แสดงได้ในรูปที่ 2.3.1



รูปที่ 2.3.1 รูปแสดงฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X

ดังนั้นเราสามารถเขียนฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X ได้ดังนี้

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{6}, & 1 \le x < 2 \\ \frac{1}{2}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

2.3.2.2 ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มแบบต่อเนื่อง

<u>นิยามที่ 2.3.6</u> ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มแบบต่อเนื่องที่ค่า ๆ หนึ่ง มีค่าเท่ากับการอินทิเกรตฟังก์ชันความน่าจะเป็นตั้งแต่ $-\infty$ จนถึงค่า ๆ นั้น นั่นคือ

$$F(x)=P(X\le x)=\int_{-\infty}^{x}f(t)dt$$

<u>ตัวอย่างที่ 2.3.7</u> กำหนดให้ X เป็นตัวแปรสุ่มแบบต่อเนื่อง ซึ่งมีฟังก์ชันความน่าจะเป็นคือ

$$f(x)=\begin{cases}\frac{1}{2}-\frac{1}{8}x, & x\in(0,4)\\ 0, & x\notin(0,4)\end{cases}$$

จงหา

(1) $F(0)$

(2) $F(1)$

(3) ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X

วิธีทำ จากนิยามฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่มแบบต่อเนื่อง

$$F(x)=\int_{-\infty}^{x}f(t)dt$$

(1) $F(0)=\int_{-\infty}^{0}f(t)dt=\int_{-\infty}^{0}0dt=0$

(2) $F(1)=\int_{-\infty}^{1}f(t)dt=\int_{-\infty}^{0}0dt+\int_{0}^{1}\left(\frac{1}{2}-\frac{1}{8}t\right)dt=\left[\frac{t}{2}-\frac{t^2}{16}\right]\Bigg|_{t=0}^{t=1}=\frac{1}{2}-\frac{1}{16}=\frac{7}{16}$

(3) ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X เป็นไปตามสมการ

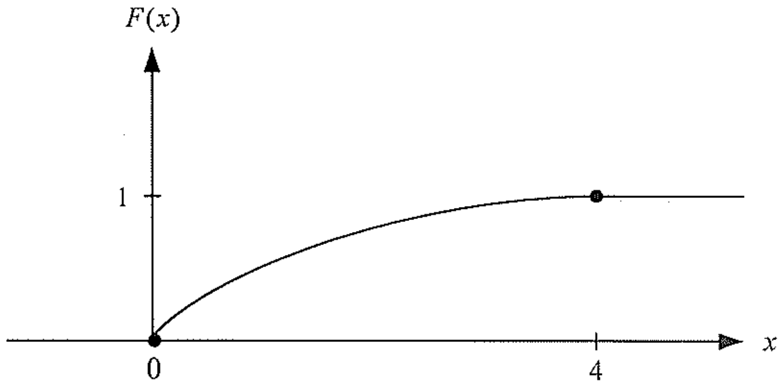
$$\begin{aligned}F(x)&=\int_{-\infty}^{x}f(t)dt\\ &=\begin{cases}\int_{-\infty}^{x}0dt, & x\le 0\\ \int_{-\infty}^{0}0dt+\int_{0}^{x}\left(\frac{1}{2}-\frac{1}{8}t\right)dt, & 0<x<4\\ \int_{-\infty}^{0}0dt+\int_{0}^{4}\left(\frac{1}{2}-\frac{1}{8}t\right)dt+\int_{0}^{x}0dt, & x\ge 4\end{cases}\end{aligned}$$

$$= \begin{cases} 0, & x \le 0 \\ \frac{1}{2}x - \frac{1}{16}x^2, & 0 < x < 4 \\ 1, & x \ge 4 \end{cases}$$

ดังนั้นฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X เป็นไปตามสมการ

$$F(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{2}x - \frac{1}{16}x^2, & 0 < x < 4 \\ 1, & x \ge 4 \end{cases}$$ ■

กราฟของฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X แสดงได้ในรูปที่ 2.3.2



รูปที่ 2.3.2 รูปแสดงฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X

ตัวอย่างที่ 2.3.8 กำหนดให้ตัวแปรสุ่ม X แทนช่วงเวลาระยะห่างในการโทรเข้าหาศูนย์วิทยุ 191 ซึ่งมีฟังก์ชันความน่าจะเป็นคือ

$$f(x) = \begin{cases} \frac{1}{20}e^{-x/20}, & x \in [0,\infty) \\ 0, & x \notin [0,\infty) \end{cases}$$

จงหาฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X

วิธีทำ เนื่องจาก X เป็นตัวแปรสุ่มแบบต่อเนื่อง ดังนั้นฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X เป็นไปตามสมการ

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

$$= \begin{cases} \int_{-\infty}^{x} 0dt, & x < 0 \\ \int_{-\infty}^{0} 0dt + \int_{0}^{x} \frac{1}{20} e^{-t/20} dt, & x \geq 0 \end{cases}$$

$$= \begin{cases} 0, & x < 0 \\ \left[-e^{-t/20} \right]_{t=0}^{t=x}, & x \geq 0 \end{cases}$$

$$= \begin{cases} 0, & x < 0 \\ 1 - e^{-x/20}, & x \geq 0 \end{cases}$$

ดังนั้นฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X เป็นไปตามสมการ

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x/20}, & x \geq 0 \end{cases}$$ ■

ความสัมพันธ์ระหว่างฟังก์ชันความน่าจะเป็นและฟังก์ชันการแจกแจงสะสม

จากหัวข้อที่ผ่านมา เราสามารถหาฟังก์ชันการแจกแจงสะสม $F(x)$ จากฟังก์ชันความน่าจะเป็น $f(x)$ ได้ ในทางกลับกัน ถ้าเราต้องการหาฟังก์ชันความน่าจะเป็น $f(x)$ จากฟังก์ชันการแจกแจงสะสม $F(x)$ จะสามารถหาได้จากความสัมพันธ์ต่อไปนี้

<u>กรณีที่ 1</u>: ถ้า X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง โดยมีค่าที่เป็นไปได้ทั้งหมดของ X คือ $x_1, x_2, \ldots, x_n$ (หรือ $x_1, x_2, \ldots$) ดังนั้น

$$f(x) = \begin{cases} F(x_1), & x = x_1 \\ F(x_i) - F(x_{i-1}), & x = x_i,\ i = 2,3,\ldots,n\ (i = 2,3,\ldots) \\ 0, & x \neq x_i,\ i = 1,2,\ldots,n\ (i = 1,2,\ldots) \end{cases}$$

<u>กรณีที่ 2</u>: ถ้า X เป็นตัวแปรสุ่มแบบต่อเนื่อง ดังนั้น

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

<u>ตัวอย่างที่ 2.3.9</u> กำหนดให้ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง โดยมีค่าที่เป็นไปได้ทั้งหมดของ X คือ 1, 2 และ 3 ถ้าตัวแปรสุ่ม X มีฟังก์ชันการแจกแจงสะสมคือ

$$F(x)=\begin{cases} 0, & x<1 \\ \frac{1}{6}, & 1\le x<2 \\ \frac{1}{2}, & 2\le x<3 \\ 1, & x\ge 3 \end{cases}$$

จงหาฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X

วิธีทำ เนื่องจาก X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ดังนั้นฟังก์ชันความน่าจะเป็นของ X คือ

$$f(x)=\begin{cases} F(1), & x=1 \\ F(2)-F(1), & x=2 \\ F(3)-F(2), & x=3 \\ 0, & x\ne 1,2,3 \end{cases}$$

$$=\begin{cases} \frac{1}{6}, & x=1 \\ \frac{1}{3}, & x=2 \\ \frac{1}{2}, & x=3 \\ 0, & x\ne 1,2,3 \end{cases}$$

ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการต่อไปนี้

$$f(x)=\begin{cases} \frac{1}{6}, & x=1 \\ \frac{1}{3}, & x=2 \\ \frac{1}{2}, & x=3 \\ 0, & x\ne 1,2,3 \end{cases}$$

■

ตัวอย่างที่ 2.3.10 กำหนดให้ตัวแปรสุ่ม X แทนช่วงเวลาระยะห่างในการโทรเข้าหาศูนย์วิทยุ 191 มีฟังก์ชันการแจกแจงสะสมคือ

$$F(x)=\begin{cases}0, & x<0\\ 1-e^{-x/20}, & x\geq 0\end{cases}$$

จงหาฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X

วิธีทำ เนื่องจาก X เป็นตัวแปรสุ่มแบบต่อเนื่อง ดังนั้นฟังก์ชันความน่าจะเป็นของ X คือ

$$\begin{aligned} f(x) &= \frac{d}{dx}F(x) \\ &= \begin{cases}\frac{d}{dx}(0), & x<0\\ \frac{d}{dx}(1-e^{-x/20}), & x\geq 0\end{cases} \\ &= \begin{cases}0, & x<0\\ \frac{1}{20}e^{-x/20}, & x\geq 0\end{cases}\end{aligned}$$

ดังนั้นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการนี้

$$f(x)=\begin{cases}0, & x<0\\ \frac{1}{20}e^{-x/20}, & x\geq 0\end{cases}$$

■

2.4 การแจกแจงความน่าจะเป็นของตัวแปรสุ่ม 2 ตัว

ในหัวข้อนี้ได้ขยายแนวความคิดของหัวข้อที่ผ่านมา นั่นคือ ขยายจากการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม 1 ตัว ไปเป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่ม 2 ตัว ดังนั้นจึงมีฟังก์ชันบางอย่างที่เพิ่มเติมขึ้นมาจากหัวข้อที่แล้ว เพื่อความสะดวกจึงได้แบ่งการพิจารณาฟังก์ชันต่าง ๆ ออกเป็นสองกรณีตามประเภทของตัวแปรสุ่มคือ กรณีที่ตัวแปรสุ่มทั้งสองตัวเป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง และกรณีที่ตัวแปรสุ่มทั้งสองตัวเป็นตัวแปรสุ่มแบบต่อเนื่อง

2.4.1 กรณีที่ตัวแปรสุ่มทั้งสองตัวเป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง

ในหัวข้อนี้ได้พิจารณาฟังก์ชันที่สำคัญของตัวแปรสุ่ม X และ Y ซึ่งเป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง

2.4.1.1 ฟังก์ชันความน่าจะเป็นร่วม

นิยามที่ 2.4.1 ***ฟังก์ชันความน่าจะเป็นร่วม (joint probability function)*** ของตัวแปรสุ่ม X และ Y แทนด้วยสัญลักษณ์ $f(x,y)$ สำหรับกรณีที่ X และ Y เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ฟังก์ชันความน่าจะเป็นร่วมมีค่าเท่ากับความน่าจะเป็นของตัวแปรสุ่ม X ที่มีค่าเท่ากับ x และตัวแปรสุ่ม Y ที่มีค่าเท่ากับ y ดังนั้น

$$f(x,y) = P(X = x, Y = y)$$

สมบัติที่สำคัญของฟังก์ชันความน่าจะเป็นร่วมสำหรับตัวแปรสุ่มแบบไม่ต่อเนื่อง

1. ฟังก์ชันความน่าจะเป็นร่วมมีค่าตั้งแต่ศูนย์ถึงหนึ่ง สำหรับทุกจำนวนจริง x และ y นั่นคือ

$$0 \le f(x,y) \le 1$$

สำหรับทุก $x, y \in \mathbb{R}$

2. ผลรวมของฟังก์ชันความน่าจะเป็นสำหรับทุกจำนวนจริง x และ y มีค่าเท่ากับหนึ่งเสมอ นั่นคือ

$$\sum_{x\in\mathbb{R}} \sum_{y\in\mathbb{R}} f(x,y) = 1$$

3. ความน่าจะเป็นของเหตุการณ์ A (ซึ่งเป็นเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม X และ Y) มีค่าเท่ากับผลรวมของฟังก์ชันความน่าจะเป็นร่วมสำหรับทุกค่าของตัวแปรสุ่ม X และ Y ที่เป็นสมาชิกของเหตุการณ์ A นั่นคือ

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

2.4.1.2 ฟังก์ชันการแจกแจงสะสมร่วม

<u>นิยามที่ 2.4.2</u> *ฟังก์ชันการแจกแจงสะสมร่วม (joint distribution function)* ของตัวแปรสุ่ม X และ Y แทนด้วยสัญลักษณ์ $F(x,y)$ สำหรับกรณีที่ X และ Y เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ฟังก์ชันการแจกแจงสะสมร่วมที่ค่า x และ y ชุดหนึ่ง จะมีค่าเท่ากับผลรวมของฟังก์ชันความน่าจะเป็นร่วมของค่า x และ y ทุกชุดที่มีค่าน้อยกว่าหรือเท่ากับค่า x และ y ชุดนั้น ๆ นั่นคือ

$$F(x,y) = P(X \le x, Y \le y) = \sum_{a\le x}\sum_{b\le y} P(X=a, Y=b) = \sum_{a\le x}\sum_{b\le y} f(a,b)$$

2.4.1.3 ฟังก์ชันความน่าจะเป็นเดี่ยว

<u>นิยามที่ 2.4.3</u> *ฟังก์ชันความน่าจะเป็นเดี่ยว (marginal probability function)*
ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X แทนด้วยสัญลักษณ์ $f_X(x)$ ซึ่งมีนิยามคือ

$$f_X(x) = P(X=x) = \sum_y f(x,y)$$

ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y แทนด้วยสัญลักษณ์ $f_Y(y)$ ซึ่งมีนิยามคือ

$$f_Y(y) = P(Y=y) = \sum_x f(x,y)$$

2.4.1.4 ฟังก์ชันการแจกแจงสะสมเดี่ยว

<u>นิยามที่ 2.4.4</u> *ฟังก์ชันการแจกแจงสะสมเดี่ยว (marginal cumulative distribution function)*
ฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม X แทนด้วยสัญลักษณ์ $F_X(x)$ ซึ่งมีนิยามคือ

$$F_X(x) = P(X \le x) = \sum_x f_X(x)$$

ฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม Y แทนด้วยสัญลักษณ์ $F_Y(y)$ ซึ่งมีนิยามคือ

$$F_Y(y) = P(Y \le y) = \sum_y f_Y(y)$$

<u>ตัวอย่างที่ 2.4.1</u> กำหนดให้ X และ Y เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มีฟังก์ชันความน่าจะเป็นร่วมดังนี้

$$f(x,y) = \begin{cases} A(x+y^2), & x=1,4 \text{ และ } y=-1,0,1,3 \\ 0, & x\ne 1,4 \text{ หรือ } y \ne -1,0,1,3 \end{cases}$$

จงหา

(1) ค่า A

(2) ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X และ Y

(3) $F(4,0)$

(4) $P(X \geq 1, 0 < Y \leq 2)$

(5) $F_X(4)$

(6) $P(X \leq 3)$

(7) $F_Y(0)$

วิธีทำ (1) จากสมบัติของฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบไม่ต่อเนื่อง จะได้ว่า

$$\sum_{x\in\mathbb{R}}\sum_{y\in\mathbb{R}} f(x,y) = 1$$

ดังนั้น

$$\sum_{x\in\mathbb{R}}\{f(x,-1)+f(x,0)+f(x,1)+f(x,3)\} = 1$$

$$\sum_{x\in\mathbb{R}} f(x,-1)+\sum_{x\in\mathbb{R}} f(x,0)+\sum_{x\in\mathbb{R}} f(x,1)+\sum_{x\in\mathbb{R}} f(x,3) = 1$$

$$\{f(1,-1)+f(4,-1)\}+\{f(1,0)+f(4,0)\}+\{f(1,1)+f(4,1)\}+\{f(1,3)+f(4,3)\} = 1$$

$$\{2A+5A\}+\{A+4A\}+\{2A+5A\}+\{10A+13A\} = 1$$

$$42A = 1$$

$$A = \frac{1}{42}$$

ดังนั้นฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่ม X และ Y เป็นดังนี้

$$f(x,y) = \begin{cases} \frac{1}{42}(x+y^2), & x = 1,4 \text{ และ } y = -1,0,1,3 \\ 0, & x \neq 1,4 \text{ หรือ } y \neq -1,0,1,3 \end{cases}$$

(2) จากนิยามของฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่มแบบไม่ต่อเนื่องจะได้ว่า

$$\begin{aligned} f_X(x) &= \sum_y f(x,y) \\ &= f(x,-1)+f(x,0)+f(x,1)+f(x,3) \\ &= \begin{cases} \frac{1}{42}(x+1)+\frac{1}{42}x+\frac{1}{42}(x+1)+\frac{1}{42}(x+9), & x = 1,4 \\ 0, & x \neq 1,4 \end{cases} \\ &= \begin{cases} \frac{1}{42}(4x+11), & x = 1,4 \\ 0, & x \neq 1,4 \end{cases} \end{aligned}$$

ดังนั้นฟังก์ชันความน่าจะเป็นเดี่ยวของ X เป็นดังนี้

$$f_X(x) = \begin{cases} \frac{1}{42}(4x+11), & x = 1,4 \\ 0, & x \neq 1,4 \end{cases}$$

จากนิยามของฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่มแบบไม่ต่อเนื่องจะได้ว่า

$$\begin{aligned} f_Y(y) &= \sum_x f(x,y) \\ &= f(1,y)+f(4,y) \\ &= \begin{cases} \frac{1}{42}(1+y^2)+\frac{1}{42}(4+y^2), & y=-1,0,1,3 \\ 0, & y\neq -1,0,1,3 \end{cases} \\ &= \begin{cases} \frac{1}{42}(5+2y^2), & y=-1,0,1,3 \\ 0, & y\neq -1,0,1,3 \end{cases} \end{aligned}$$

ดังนั้นฟังก์ชันความน่าจะเป็นเดี่ยวของ Y เป็นดังนี้

$$f_Y(y)=\begin{cases} \frac{1}{42}(5+2y^2), & y=-1,0,1,3 \\ 0, & y\neq -1,0,1,3 \end{cases}$$

(3) เมื่อทราบฟังก์ชันความน่าจะเป็นร่วม สามารถหาฟังก์ชันการแจกแจงสะสมร่วมได้ดังนี้

$$\begin{aligned} F(4,0) &= \sum_{a\le 4}\sum_{b\le 0} f(a,b) \\ &= \sum_{a\le 4}\{f(a,-1)+f(a,0)\} \\ &= \{f(1,-1)+f(1,0)\}+\{f(4,-1)+f(4,0)\} \\ &= \left(\frac{2}{42}+\frac{1}{42}\right)+\left(\frac{5}{42}+\frac{4}{42}\right)=\frac{12}{42}=\frac{6}{21} \end{aligned}$$

ดังนั้นฟังก์ชันการแจกแจงสะสมร่วมเมื่อ $x=4$ และ $y=0$ มีค่าเท่ากับ $\frac{6}{21}$

(4) เมื่อทราบฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบไม่ต่อเนื่อง สามารถหาค่าความน่าจะเป็นที่ต้องการได้ดังนี้

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

พิจารณาเหตุการณ์ $A=\{X\ge 1, 0<Y\le 2\}$

$$P(X\ge 1,0<Y\le 2)=\sum_{x\ge 1}\sum_{0<y\le 2} f(x,y)=\sum_{x\ge 1} f(x,1)=f(1,1)+f(4,1)=\frac{2}{42}+\frac{5}{42}=\frac{1}{6}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์เมื่อ $x\ge 1$ และ $0<y\le 2$ มีค่าเท่ากับ $\frac{1}{6}$

(5) จากนิยามฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม X และฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X ที่หามาแล้วจากข้อ (2) ทำให้หาค่า $F_X(4)$ ได้ดังนี้

$$F_X(4) = P(X \le 4) = \sum_{a \le 4} f_X(a) = f_X(1) + f_X(4) = \frac{1}{42}(4+11) + \frac{1}{42}(16+11) = 1$$

ดังนั้นฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม X เมื่อ $x = 4$ มีค่าเท่ากับ 1

(6) จากฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X ที่หามาแล้วจากข้อ (2) และใช้การพิจารณาเหมือนมีตัวแปรสุ่ม X เพียงตัวเดียว เช่นเดียวกับในหัวข้อที่ 2.3 สามารถหาค่า $P(X \le 3)$ ได้ดังนี้

$$P(X \le 3) = \sum_{a \le 3} f_X(a) = f_X(1) = \frac{15}{42} = \frac{5}{14}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์ที่ตัวแปรสุ่ม X เมื่อ $x \le 3$ มีค่าเท่ากับ $\frac{5}{14}$

(7) จากนิยามของฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม Y และฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y ที่หามาแล้วจากข้อ (2) ทำให้หาค่า $F_Y(0)$ ได้ดังนี้

$$F_Y(0) = P(Y \le 0) = \sum_{a \le 0} f_Y(a) = f_Y(-1) + f_Y(0) = \frac{1}{42}(5+2) + \frac{1}{42}(5+0) = \frac{12}{42} = \frac{2}{7}$$

ดังนั้นฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม Y เมื่อ $y = 0$ มีค่าเท่ากับ $\frac{2}{7}$ ∎

2.4.2 กรณีที่ตัวแปรสุ่มทั้งสองเป็นตัวแปรสุ่มแบบต่อเนื่อง

ในหัวข้อนี้ได้พิจารณาฟังก์ชันที่สำคัญของตัวแปรสุ่ม X และ Y ซึ่งเป็นตัวแปรสุ่มแบบต่อเนื่อง

2.4.2.1 ฟังก์ชันความน่าจะเป็นร่วม

<u>นิยามที่ 2.4.5</u> กำหนดให้ X และ Y เป็นตัวแปรสุ่มแบบต่อเนื่อง พิจารณาฟังก์ชัน f ที่ส่งจากเซตของจำนวนจริงไปยังเซตของจำนวนจริง โดยที่

$$f(x,y) \ge 0$$

สำหรับทุกค่า $x, y \in \mathbb{R}$ และ

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

เรียกฟังก์ชัน f ซึ่งสอดคล้องกับเงื่อนไขเหล่านี้ว่า ***ฟังก์ชันความน่าจะเป็นร่วม (joint probability function)*** ของตัวแปรสุ่ม X และ Y

สมบัติที่สำคัญของฟังก์ชันความน่าจะเป็นร่วมสำหรับตัวแปรสุ่มแบบต่อเนื่อง

1. ฟังก์ชันความน่าจะเป็นร่วมมีค่ามากกว่าหรือเท่ากับศูนย์สำหรับทุกจำนวนจริง x และ y นั่นคือ

$$f(x,y) \ge 0$$

สำหรับทุกค่า $x, y \in \mathbb{R}$

2. อินทิเกรตฟังก์ชันความน่าจะเป็นร่วมเมื่อ x และ y อยู่ในช่วง $(-\infty,\infty)$ มีค่าเท่ากับหนึ่งเสมอนั่นคือ

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$$

3. ความน่าจะเป็นของเหตุการณ์ $A=\{a \le X \le b, c \le Y \le d\}$ มีค่าเท่ากับอินทิเกรตฟังก์ชันความน่าจะเป็นเมื่อ x อยู่ในช่วง (a,b) และ y อยู่ในช่วง (c,d) นั่นคือ

$$P(a \le X \le b, c \le Y \le d) = \int_{c}^{d}\int_{a}^{b} f(x,y)dxdy$$

โดยที่ $-\infty < a \le b < \infty$ และ $-\infty < c \le d < \infty$

2.4.2.2 ฟังก์ชันการแจกแจงสะสมร่วม

<u>นิยามที่ 2.4.6</u> *ฟังก์ชันการแจกแจงสะสมร่วม (joint distribution function)* ของตัวแปรสุ่มแบบต่อเนื่อง X และ Y แทนด้วยสัญลักษณ์ $F(x,y)$ และมีนิยามดังนี้

$$F(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv$$

2.4.2.3 ฟังก์ชันความน่าจะเป็นเดี่ยว

<u>นิยามที่ 2.4.7</u> *ฟังก์ชันความน่าจะเป็นเดี่ยว (marginal probability function)*

ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X แทนด้วยสัญลักษณ์ $f_X(x)$ ซึ่งมีนิยามคือ

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y แทนด้วยสัญลักษณ์ $f_Y(y)$ ซึ่งมีนิยามคือ

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

2.4.2.4 ฟังก์ชันการแจกแจงสะสมเดี่ยว

<u>นิยามที่ 2.4.8</u> *ฟังก์ชันการแจกแจงสะสมเดี่ยว (marginal cumulative distribution function)*

ฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม X แทนด้วยสัญลักษณ์ $F_X(x)$ ซึ่งมีนิยามคือ

$$F_X(x) = \int_{-\infty}^{\infty} f_X(x)dx$$

ฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม Y แทนด้วยสัญลักษณ์ $F_Y(y)$ ซึ่งมีนิยามคือ

$$F_Y(x) = \int_{-\infty}^{\infty} f_Y(y)dy$$

ตัวอย่างที่ 2.4.2 กำหนดให้ X และ Y เป็นตัวแปรสุ่มแบบต่อเนื่องที่มีฟังก์ชันความน่าจะเป็นร่วมดังนี้

$$f(x,y)=\begin{cases} kx^2y, & x\in(0,1) \text{ และ } y\in(0,1) \\ 0, & x\notin(0,1) \text{ หรือ } y\notin(0,1) \end{cases}$$

จงหา

(1) ค่า k

(2) ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X และ Y

(3) $F(\frac{1}{2},\frac{1}{4})$

(4) $P(0.5<X\le 1, 0<Y<1)$

วิธีทำ (1) จากสมบัติของฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบต่อเนื่อง X และ Y จะได้ว่า

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy=1$$

ดังนั้น

$$\int_0^1\int_0^1 kx^2ydxdy=1$$

$$\int_0^1\left[\frac{kx^3y}{3}\right]_{x=0}^{x=1} dy=1$$

$$\int_0^1\frac{ky}{3}dy=1$$

$$\left[\frac{ky^2}{6}\right]_{y=0}^{y=1} dy=1$$

$$\frac{k}{6}=1$$

$$k=6$$

(2) จากนิยามของฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่มแบบต่อเนื่องจึงได้ว่า

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy$$

$$=\begin{cases} \int_0^1 6x^2ydy, & x\in(0,1) \\ 0, & x\notin(0,1) \end{cases}$$

$$=\begin{cases} \left[3x^2y^2\right]_{y=0}^{y=1}, & x\in(0,1) \\ 0, & x\notin(0,1) \end{cases}$$

$$= \begin{cases} 3x^2, & x \in (0,1) \\ 0, & x \notin (0,1) \end{cases}$$

ดังนั้นฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X เป็นดังนี้

$$f_X(x) = \begin{cases} 3x^2, & x \in (0,1) \\ 0, & x \notin (0,1) \end{cases}$$

จากนิยามของฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่มแบบต่อเนื่องจึงได้ว่า

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

$$= \begin{cases} \int_0^1 6x^2 y dx, & y \in (0,1) \\ 0, & y \notin (0,1) \end{cases}$$

$$= \begin{cases} \left[2x^3 y\right]_{x=0}^{x=1}, & y \in (0,1) \\ 0, & y \notin (0,1) \end{cases}$$

$$= \begin{cases} 2y, & y \in (0,1) \\ 0, & y \notin (0,1) \end{cases}$$

ดังนั้นฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y เป็นดังนี้

$$f_Y(x) = \begin{cases} 2y, & y \in (0,1) \\ 0, & y \notin (0,1) \end{cases}$$

(3) จากนิยามของฟังก์ชันการแจกแจงสะสมร่วมของตัวแปรสุ่มแบบต่อเนื่องจึงได้ว่า

$$F\left(\frac{1}{2},\frac{1}{4}\right) = \int_{-\infty}^{1/4}\int_{-\infty}^{1/2} f(x,y)dxdy = \int_0^{1/4}\int_0^{1/2} 6x^2 y dxdy$$

$$= \int_0^{1/4} \left[2x^3 y\right]_{x=0}^{x=1/2} dy = \int_0^{1/4} \frac{y}{4} dy = \left[\frac{y^2}{8}\right]_{y=0}^{y=1/4} = \frac{1}{128}$$

ดังนั้นฟังก์ชันการแจกแจงสะสมร่วมเมื่อ $x = \frac{1}{2}$ และ $y = \frac{1}{4}$ มีค่าเท่ากับ $\frac{1}{128}$

(4) จากสมบัติของฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$P(0.5 < X \le 1, 0 < Y < 1) = \int_0^1 \int_{0.5}^1 6x^2 y dxdy = \int_0^1 \left[2x^3 y\right]_{x=0.5}^{x=1} dy$$

$$= \int_0^1 \frac{7y}{4} dy = \left[\frac{7y^2}{8}\right]_{y=0}^{y=1} = \frac{7}{8}$$

ดังนั้นความน่าจะเป็นของเหตุการณ์เมื่อ $0.5 < x \le 1$ และ $0 < y < 1$ มีค่าเท่ากับ $\frac{7}{8}$ ■

<u>**ตัวอย่างที่ 2.4.3**</u> กำหนดให้ฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบต่อเนื่อง X และ Y เป็นดังนี้

$$f(x,y)=\begin{cases}4xy, & x\in(0,1) \text{ และ } y\in(0,1)\\ 0, & x\notin(0,1) \text{ หรือ } y\notin(0,1)\end{cases}$$

จงหา

(1) ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X และ Y

(2) $F_X(0.5)$

(3) $P(Y\le 0.5)$

(4) $P(X\ge 0, Y\le 0.5)$

วิธีทำ (1) จากนิยามของฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่มแบบต่อเนื่องจึงได้ว่า

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\ &= \begin{cases}\int_0^1 4xydy, & x\in(0,1)\\ 0, & x\notin(0,1)\end{cases} \\ &= \begin{cases}\left[2xy^2\right]_{y=0}^{y=1}, & x\in(0,1)\\ 0, & x\notin(0,1)\end{cases} \\ &= \begin{cases}2x, & x\in(0,1)\\ 0, & x\notin(0,1)\end{cases} \end{aligned}$$

ดังนั้นฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X เป็นดังนี้

$$f_X(x)=\begin{cases}2x, & x\in(0,1)\\ 0, & x\notin(0,1)\end{cases}$$

จากนิยามของฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่มแบบต่อเนื่องจึงได้ว่า

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx \\ &= \begin{cases}\int_0^1 4xydx, & y\in(0,1)\\ 0, & y\notin(0,1)\end{cases} \\ &= \begin{cases}\left[2x^2y\right]_{x=0}^{x=1}, & y\in(0,1)\\ 0, & y\notin(0,1)\end{cases} \\ &= \begin{cases}2y, & y\in(0,1)\\ 0, & y\notin(0,1)\end{cases} \end{aligned}$$

ดังนั้นฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y เป็นดังนี้

$$f_Y(x) = \begin{cases} 2y, & y \in (0,1) \\ 0, & y \notin (0,1) \end{cases}$$

(2) จากนิยามฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม X และฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X ที่หามาแล้วจากข้อ (1) ทำให้หาค่า $F_X(0.5)$ ได้ดังนี้

$$F_X(0.5) = \int_{-\infty}^{0.5} f_X(x)dx = \int_0^{0.5} 2xdx = \left[x^2\right]_{x=0}^{x=0.5} = \frac{1}{4} = 0.25$$

ดังนั้นฟังก์ชันการแจกแจงสะสมเดี่ยวของตัวแปรสุ่ม X เมื่อ $x = 0.5$ มีค่าเท่ากับ 0.25

(3) จากฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y ที่หามาแล้วจากข้อ (1) และใช้การพิจารณาเหมือนมีตัวแปรสุ่ม Y เพียงตัวเดียว เช่นเดียวกับในหัวข้อที่ 2.3 สามารถหาค่า $P(Y \le 0.5)$ ได้ดังนี้

$$P(Y \le 0.5) = \int_{-\infty}^{0.5} f_Y(y)dy = \int_0^{0.5} 2ydy = \left[y^2\right]_{y=0}^{y=0.5} = \frac{1}{4} = 0.25$$

ดังนั้นความน่าจะเป็นของเหตุการณ์ที่ตัวแปรสุ่ม Y เมื่อ $y \le 0.5$ มีค่าเท่ากับ 0.25

(4) จากสมบัติของฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$P(X \ge 0, Y \le 0.5) = \int_{-\infty}^{0.5}\int_0^{\infty} 4xydxdy = \int_0^{0.5}\int_0^{1} 4xydxdy$$

$$= \int_0^{0.5}\left[2x^2y\right]_{x=0}^{x=1} dy = \int_0^{0.5} 2ydy = \left[y^2\right]_{y=0}^{y=0.5} = 0.25$$

ดังนั้นความน่าจะเป็นของเหตุการณ์เมื่อ $x \ge 0$ และ $y \le 0.5$ มีค่าเท่ากับ 0.25 ■

2.5 ความเป็นอิสระต่อกันของตัวแปรสุ่ม 2 ตัว

<u>นิยามที่ 2.5.1</u> กำหนดให้ X และ Y เป็นตัวแปรสุ่มใด ๆ $f(x,y)$, $f_X(x)$ และ $f_Y(y)$ แทนฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่ม X, Y ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X และฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y ตามลำดับ ดังนั้นตัวแปรสุ่ม X และ Y จะเป็นอิสระต่อกันถ้า

$$f(x,y) = f_X(x)f_Y(y)$$

สำหรับทุกจำนวนจริง x และ y

<u>ตัวอย่างที่ 2.5.1</u> กำหนดให้ X และ Y เป็นตัวแปรสุ่มแบบไม่ต่อเนื่องที่มีฟังก์ชันความน่าจะเป็นร่วมตามตาราง ต่อไปนี้

X \ Y	-3	2	4
1	0.1	0.2	0.2
3	0.3	0.1	0.1

จงตรวจสอบว่าตัวแปรสุ่ม X และ Y เป็นอิสระต่อกันหรือไม่

วิธีทำ จากตาราง

$$f(1,-3)=0.1$$

พิจารณา

$$f_X(1)=\sum_y f(1,y)=f(1,-3)+f(1,2)+f(1,4)=0.1+0.2+0.2=0.5$$

และ

$$f_y(-3)=\sum_x f(x,-3)=f(1,-3)+f(3,-3)=0.1+0.3=0.4$$

ดังนั้น

$$f_X(1)f_Y(-3)=(0.5)(0.4)=0.2$$

ซึ่งทำให้

$$f(1,-3)\neq f_X(1)f_Y(-3)$$

ดังนั้นตัวแปรสุ่ม X และ Y จึงไม่เป็นอิสระต่อกัน ■

<u>ตัวอย่างที่ 2.5.2</u> กำหนดฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบต่อเนื่อง X และ Y เป็นดังนี้

$$f(x,y)=\begin{cases}4xy, & x\in(0,1) \text{ และ } y\in(0,1)\\ 0, & x\notin(0,1) \text{ หรือ } y\notin(0,1)\end{cases}$$

จงตรวจสอบความเป็นอิสระของตัวแปรสุ่ม X และ Y

วิธีทำ จากตัวอย่างที่ 2.4.3 ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X เป็นดังนี้

$$f_X(x)=\begin{cases}2x, & x\in(0,1)\\ 0, & x\notin(0,1)\end{cases}$$

และฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y เป็นดังนี้

$$f_Y(x)=\begin{cases}2y, & y\in(0,1)\\ 0, & y\notin(0,1)\end{cases}$$

ดังนั้น

$$f_X(x)f_Y(y)=\begin{cases}4xy, & x\in(0,1) \text{ และ } y\in(0,1)\\ 0, & x\notin(0,1) \text{ หรือ } y\notin(0,1)\end{cases}$$

นั่นคือ

$$f(x,y)=f_X(x)f_Y(y)$$

สำหรับทุกค่า x และ y ดังนั้นจึงสามารถสรุปได้ว่า ตัวแปรสุ่ม X และ Y เป็นอิสระต่อกัน ■

<u>ตัวอย่างที่ 2.5.3</u> กำหนดฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มแบบต่อเนื่อง X และ Y เป็นดังนี้

$$f(x,y)=\begin{cases} 8e^{-(2x+4y)}, & x\in(0,\infty) \text{ และ } y\in(0,\infty) \\ 0, & x\notin(0,\infty) \text{ หรือ } y\notin(0,\infty) \end{cases}$$

จงตรวจสอบความเป็นอิสระของตัวแปรสุ่ม X และ Y

วิธีทำ ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม X หาได้ดังนี้

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\
&= \begin{cases} \int_0^{\infty} 8e^{-(2x+4y)}dy, & x\in(0,\infty) \\ 0, & x\notin(0,\infty) \end{cases} \\
&= \begin{cases} 8e^{-2x}\int_0^{\infty} e^{-4y}dy, & x\in(0,\infty) \\ 0, & x\notin(0,\infty) \end{cases} \\
&= \begin{cases} 8e^{-2x}\left[-\frac{1}{4}e^{-4y}\right]_{y=0}^{y=\infty}, & x\in(0,\infty) \\ 0, & x\notin(0,\infty) \end{cases} \\
&= \begin{cases} 2e^{-2x}, & x\in(0,\infty) \\ 0, & x\notin(0,\infty) \end{cases}
\end{aligned}$$

ฟังก์ชันความน่าจะเป็นเดี่ยวของตัวแปรสุ่ม Y หาได้ดังนี้

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx \\
&= \begin{cases} \int_0^{\infty} 8e^{-(2x+4y)}dx, & y\in(0,\infty) \\ 0, & y\notin(0,\infty) \end{cases} \\
&= \begin{cases} 8e^{-4y}\int_0^{\infty} e^{-2x}dx, & y\in(0,\infty) \\ 0, & y\notin(0,\infty) \end{cases}
\end{aligned}$$

$$=\begin{cases} 8e^{-4y}\left[-\frac{1}{2}e^{-2x}\right]_{x=0}^{x=\infty}, & y\in(0,\infty) \\ 0, & y\notin(0,\infty) \end{cases}$$
$$=\begin{cases} 4e^{-4y}, & y\in(0,\infty) \\ 0, & y\notin(0,\infty) \end{cases}$$

ดังนั้น

$$f_X(x)f_Y(y)=\begin{cases} 8e^{-(2x+4y)}, & x\in(0,\infty) \text{ และ } y\in(0,\infty) \\ 0, & x\notin(0,\infty) \text{ หรือ } y\notin(0,\infty) \end{cases}$$

นั่นคือ

$$f(x,y)=f_X(x)f_Y(y)$$

สำหรับทุกค่า x และ y ดังนั้นจึงสามารถสรุปได้ว่า ตัวแปรสุ่ม X และ Y เป็นอิสระต่อกัน ■

2.6 ค่าคาดหมาย

ค่าคาดหมาย (expected value) ของตัวแปรสุ่ม X เขียนแทนด้วยสัญลักษณ์ $E(X)$ แยกพิจารณาออกเป็นค่าคาดหมายของตัวแปรสุ่มแบบไม่ต่อเนื่อง และค่าคาดหมายของตัวแปรสุ่มแบบต่อเนื่องดังนี้

2.6.1 ค่าคาดหมายของตัวแปรสุ่มแบบไม่ต่อเนื่อง

<u>นิยามที่ 2.6.1</u> กำหนดให้ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง และ $f(x)$ แทนฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ดังนั้นสามารถนิยามค่าคาดหมายของตัวแปรสุ่ม X ได้ตามสมการต่อไปนี้

$$E(X)=\sum_x xf(x)$$

สำหรับกรณีที่ต้องการหาค่าคาดหมายของฟังก์ชัน $h(x)$ ที่ขึ้นกับตัวแปรสุ่ม X ดังนั้น

$$E(h(X))=\sum_x h(x)f(x)$$

สำหรับกรณีที่ต้องการหาค่าคาดหมายของฟังก์ชัน $g(x,y)$ ที่ขึ้นอยู่กับตัวแปรสุ่ม X,Y และมีฟังก์ชันความน่าจะเป็นร่วมคือ $f(x,y)$ ดังนั้น

$$E(g(X,Y))=\sum_x\sum_y g(x,y)f(x,y)$$

2.6.2 ค่าคาดหมายของตัวแปรสุ่มแบบต่อเนื่อง

<u>นิยามที่ 2.6.2</u> กำหนดให้ X เป็นตัวแปรสุ่มแบบต่อเนื่อง และ $f(x)$ แทนฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ดังนั้นสามารถนิยามค่าคาดหมายของตัวแปรสุ่ม X ได้ตามสมการต่อไปนี้

$$E(X)=\int_{-\infty}^{\infty} xf(x)dx$$

สำหรับกรณีที่ต้องการหาค่าคาดหมายของฟังก์ชัน $h(x)$ ที่ขึ้นกับตัวแปรสุ่ม X ดังนั้น

$$E(h(X))=\int_{-\infty}^{\infty} h(x)f(x)dx$$

สำหรับกรณีที่ต้องการหาค่าคาดหมายของฟังก์ชัน $g(x,y)$ ที่ขึ้นอยู่กับตัวแปรสุ่ม X,Y และมีฟังก์ชันความน่าจะเป็นร่วมคือ $f(x,y)$ ดังนั้น

$$E(g(X,Y))=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy$$

<u>ตัวอย่างที่ 2.6.1</u> กำหนดให้ตัวแปรสุ่มแบบไม่ต่อเนื่อง X มีค่าความน่าจะเป็นตามตารางต่อไปนี้

X	0	8	12	16
$P(X=x)$	0.3	0.2	0.3	0.2

จงหา $E(X)$ และ $E\left((X-E(X))^2\right)$

วิธีทำ จากนิยามค่าคาดหมายของตัวแปรสุ่มแบบไม่ต่อเนื่อง หาค่า $E(X)$ ได้ดังนี้

$$E(X)=\sum_x xf(x)=0f(0)+8f(8)+12f(12)+16f(16)$$
$$=0(0.3)+8(0.2)+12(0.3)+16(0.2)=8.4$$

จากนิยามค่าคาดหมายของฟังก์ชันที่เกี่ยวข้องกับตัวแปรสุ่มแบบไม่ต่อเนื่อง และแทนค่า $E(X)=8.4$ จะได้ว่า

$$E\left((X-E(X))^2\right)=E\left((X-8.4)^2\right)$$
$$=\sum_x (x-8.4)^2 f(x)$$
$$=(0-8.4)^2 f(0)+(8-8.4)^2 f(8)+(12-8.4)^2 f(12)+(16-8.4)^2 f(16)$$
$$=(8.4)^2(0.3)+(0.4)^2(0.2)+(3.6)^2(0.3)+(7.6)^2(0.2)=36.64$$

ดังนั้นค่าคาดหมายของตัวแปรสุ่ม X และ $(X-E(X))^2$ มีค่าเท่ากับ 8.4 และ 36.64 ตามลำดับ ■

ตัวอย่างที่ 2.6.2 นายสมชายมีเงินอยู่ 100,000 บาทต้องการนำไปลงทุนมี 2 ทางเลือกดังนี้

1. ลงทุนซื้อพันธบัตรที่ให้อัตราดอกเบี้ยคงที่ 12%
2. ลงทุนซื้อหุ้นบริษัท A ซึ่งมีรายงานอัตราเงินปันผลในอดีต ดังตารางต่อไปนี้

อัตราเงินปันผล	ความน่าจะเป็น
30%	0.20
25%	0.20
20%	0.30
15%	0.10
10%	0.10
5%	0.10

ถามว่านายสมชายควรเลือกลงทุนแบบใด

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนอัตราเงินปันผลในการซื้อหุ้นบริษัท A ซึ่งเป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง พิจารณาค่าคาดหมายของอัตราเงินปันผล ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} E(X) &= \sum_{x} x f(x) \\ &= 30f(30)+25f(25)+20f(20)+15f(15)+10f(10)+5f(5) \\ &= 30(0.2)+25(0.2)+20(0.3)+15(0.1)+10(0.1)+5(0.1) \\ &= 20 \end{aligned}$$

จากการคำนวณคาดว่าจะได้เงินปันผลจากหุ้น A เท่ากับ 20% เมื่อเปรียบเทียบอัตราผลตอบแทนที่คาดว่าจะได้จากการซื้อหุ้นบริษัท A และอัตราผลตอบแทนที่ได้จากการซื้อพันธบัตรพบว่า อัตราผลตอบแทนที่คาดว่าจะได้จากการซื้อหุ้นบริษัท A สูงกว่าอัตราผลตอบแทนที่ได้จากการซื้อพันธบัตร ดังนั้นนายสมชายควรจะเลือกลงทุนด้วยการซื้อหุ้นบริษัท A ■

หมายเหตุ จากตัวอย่างที่ 2.6.2 พิจารณาเพียงแค่อัตราเงินปันผลที่คาดว่าจะได้รับเท่านั้น แต่ตามหลักความเป็นจริงแล้วการพิจารณาเพียงแค่ค่าคาดหมายของอัตราเงินปันผลยังไม่เพียงพอ จำเป็นจะต้องพิจารณาถึงความเสี่ยงของการลงทุนด้วย

ตัวอย่างที่ 2.6.3 ในการเล่นเกม ๆ หนึ่ง โดยการทอดลูกเต๋าเที่ยงตรง 3 ลูก ผู้เล่นจะวางเงินในการเล่นครั้งละ 10 บาท ในการทอดลูกเต๋าแต่ละครั้ง ถ้าลูกเต๋าขึ้นแต้ม 5 หนึ่งลูก ผู้เล่นจะได้รับเงิน 10 บาท ถ้าลูกเต๋าขึ้นแต้ม 5 สองลูก ผู้เล่นจะได้รับเงิน 20 บาท ถ้าลูกเต๋าขึ้นแต้ม 5 สามลูก ผู้เล่นจะได้รับเงิน 30 บาท นอกจากนั้นผู้เล่นจะไม่ได้รับเงินตอบแทนในการเล่นเกมแต่จะเสียเงินที่วางเอาไว้ 10 บาท ผู้เล่นคาดว่าจะได้รับเงินเท่าใด และในการเล่นเกมนี้ยุติธรรมหรือไม่

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนเงินที่ผู้เล่นเกมจะได้รับ นั่นคือตัวแปรสุ่ม X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ซึ่งมีค่าที่เป็นไปได้ทั้งหมดคือ –10, 10, 20 และ 30 บาท และสามารถหาค่าความน่าจะเป็นของแต่ละค่าที่เป็นไปได้ของตัวแปรสุ่ม X ได้ดังนี้

$P(X=-10)$ มีค่าเท่ากับความน่าจะเป็นของเหตุการณ์ที่ทอดลูกเต๋าไม่ขึ้นแต้มเลย ดังนั้น

$$P(X=-10)=\frac{125}{216}$$

$P(X=10)$ มีค่าเท่ากับความน่าจะเป็นของเหตุการณ์ที่ทอดลูกเต๋าขึ้นแต้ม 5 หนึ่งลูก ดังนั้น

$$P(X=10)=\frac{75}{216}$$

$P(X=20)$ มีค่าเท่ากับความน่าจะเป็นของเหตุการณ์ที่ทอดลูกเต๋าขึ้นแต้ม 5 สองลูก ดังนั้น

$$P(X=20)=\frac{15}{216}$$

$P(X=30)$ มีค่าเท่ากับความน่าจะเป็นของเหตุการณ์ที่ทอดลูกเต๋าขึ้นแต้ม 5 สามลูก ดังนั้น

$$P(X=30)=\frac{1}{216}$$

ค่าคาดหมายของจำนวนเงินที่ผู้เล่นเกมจะได้รับ หาค่าได้ดังนี้

$$\begin{aligned} E(X) &= \sum_x xf(x) \\ &= (-10)f(-10)+10f(10)+20f(20)+30f(30) \\ &= (-10)\left(\frac{125}{216}\right)+10\left(\frac{75}{216}\right)+20\left(\frac{15}{216}\right)+30\left(\frac{1}{216}\right) \\ &= -\frac{170}{216}\approx -0.79 \end{aligned}$$

เนื่องจากเกมที่ถือว่ายุติธรรมจะต้องมีค่าคาดหมายเท่ากับศูนย์ ดังนั้น เกมนี้ถือว่าไม่ยุติธรรม เพราะผู้เล่นคาดว่าจะเสียเงินจำนวน 0.79 บาท ■

<u>ตัวอย่างที่ 2.6.4</u> กำหนดให้ X เป็นตัวแปรสุ่มแบบต่อเนื่องที่มีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x)=\begin{cases} 2x, & x\in(0,1) \\ 0, & x\notin(0,1) \end{cases}$$

จงหา $E(X)$ และ $E(X^2)$

วิธีทำ จากนิยามค่าคาดหมายของตัวแปรสุ่มแบบต่อเนื่อง หาค่า $E(X)$ ได้ดังนี้

$$E(X)=\int_{-\infty}^{\infty} xf(x)dx=\int_{-\infty}^{0} x\cdot 0dx+\int_{0}^{1} x\cdot 2xdx+\int_{1}^{\infty} x\cdot 0dx=\int_{0}^{1} 2x^2dx=2\left[\frac{x^3}{3}\right]_{x=0}^{1}=\frac{2}{3}$$

จากนิยามค่าคาดหมายของฟังก์ชันที่เกี่ยวข้องกับตัวแปรสุ่มแบบต่อเนื่อง หาค่า $E(X^2)$ ได้ดังนี้

$$E(X^2)=\int_{-\infty}^{\infty}x^2 f(x)dx=\int_{-\infty}^{0}x^2\cdot 0dx+\int_{0}^{1}x^2\cdot 2xdx+\int_{1}^{\infty}x^2\cdot 0dx=\int_{0}^{1}2x^3dx=2\left[\frac{x^4}{4}\right]_{x=0}^{1}=\frac{1}{2}$$

ดังนั้นค่าคาดหมายของตัวแปรสุ่ม X และ X^2 มีค่าเท่ากับ $\frac{2}{3}$ และ $\frac{1}{2}$ ตามลำดับ ▩

ตัวอย่างที่ 2.6.5 กำหนดให้ X เป็นตัวแปรสุ่มแบบต่อเนื่องที่มีฟังก์ชันการแจกแจงสะสมดังนี้

$$F(x)=\begin{cases}0, & x<-1\\ \frac{x^3}{2}+\frac{1}{2}, & -1<x<1\\ 1, & x>1\end{cases}$$

จงหา $E(X)$

วิธีทำ เนื่องจาก X เป็นตัวแปรสุ่มแบบต่อเนื่อง สามารถหาค่าคาดหมายของตัวแปรสุ่ม X ได้ดังนี้

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx$$

แต่จากโจทย์ให้ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X มา ดังนั้นเราจำเป็นจะต้องหาฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X ก่อนดังนี้

$$f(x)=\frac{d}{dx}F(x)=\begin{cases}0, & x<-1\\ \frac{3x^2}{2}, & -1<x<1\\ 0, & x>1\end{cases}$$

โดยใช้นิยามค่าคาดหมายของตัวแปรสุ่มจึงได้ว่า

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_{-\infty}^{-1}x\cdot 0dx+\int_{-1}^{1}x\cdot\frac{3x^2}{2}dx+\int_{1}^{\infty}x\cdot 0dx=\int_{-1}^{1}\frac{3x^3}{2}dx=\frac{3}{2}\left[\frac{x^4}{4}\right]_{x=-1}^{1}=0$$

ดังนั้นค่าคาดหมายของตัวแปรสุ่ม X มีค่าเท่ากับ 0 ▩

กฎของค่าคาดหมาย

1. ค่าคาดหมายของค่าคงที่มีค่าเท่ากับค่าคงที่นั้น ๆ นั่นคือ

$$E(a)=a$$

สำหรับทุก $a\in\mathbb{R}$

2. ค่าคาดหมายของค่าคงที่คูณกับตัวแปรสุ่มมีค่าเท่ากับค่าคงที่คูณกับค่าคาดหมายของตัวแปรสุ่ม

$$E(aX)=aE(X)$$

เมื่อ $a\in\mathbb{R}$ และ X เป็นตัวแปรสุ่มใด ๆ

<u>พิสูจน์</u> กำหนดให้ X เป็นตัวแปรสุ่มใด ๆ และ $a\in\mathbb{R}$

กรณีที่ 1 ตัวแปรสุ่มแบบไม่ต่อเนื่อง

$$E(aX) = \sum_{x} axf(x) = a\sum_{x} xf(x) = aE(X)$$

กรณีที่ 2 ตัวแปรสุ่มแบบต่อเนื่อง

$$E(aX) = \int_{\infty}^{\infty} axf(x)dx = a\int_{\infty}^{\infty} xf(x)dx = aE(X)$$

3. ค่าคาดหมายมีสมบัติความเป็นเชิงเส้น นั่นคือ

$$E(aX \pm bY) = aE(X) \pm bE(Y)$$

เมื่อ $a,b \in \mathbb{R}$ และ X,Y เป็นตัวแปรสุ่มใด ๆ

<u>พิสูจน์</u> กำหนดให้ X,Y เป็นตัวแปรสุ่มใด ๆ และ $a,b \in \mathbb{R}$

กรณีที่ 1 ตัวแปรสุ่มแบบไม่ต่อเนื่อง

$$\begin{aligned}
E(aX \pm bY) &= \sum_{x}\sum_{y}(ax \pm by)f(x,y) \\
&= \sum_{x}\sum_{y} axf(x,y) \pm \sum_{x}\sum_{y} byf(x,y) \\
&= \sum_{x} ax\sum_{y} f(x,y) \pm \sum_{x} by\sum_{y} f(x,y) \\
&= a\sum_{x} xf_X(x) \pm b\sum_{x} yf_Y(y) \\
&= aE(X) \pm bE(Y)
\end{aligned}$$

กรณีที่ 2 ตัวแปรสุ่มแบบต่อเนื่อง

$$\begin{aligned}
E(aX \pm bY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax \pm by)f(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} axf(x,y)dxdy \pm \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} byf(x,y)dxdy \\
&= \int_{-\infty}^{\infty} ax\left[\int_{-\infty}^{\infty} f(x,y)dy\right]dx \pm \int_{-\infty}^{\infty} by\left[\int_{-\infty}^{\infty} f(x,y)dx\right]dy \\
&= a\int_{-\infty}^{\infty} xf_X(x)dx \pm b\int_{-\infty}^{\infty} yf_Y(y)dy \\
&= aE(X) \pm bE(Y)
\end{aligned}$$

4. ถ้า X และ Y เป็นตัวแปรสุ่มใด ๆ ที่เป็นอิสระต่อกันแล้ว

$$E(XY) = E(X)E(Y)$$

<u>พิสูจน์</u> กำหนดให้ X และ Y เป็นตัวแปรสุ่มใด ๆ ที่เป็นอิสระต่อกัน

กรณีที่ 1 ตัวแปรสุ่มแบบไม่ต่อเนื่อง

$$E(XY) = \sum_{x}\sum_{y} xyf(x,y)$$

เนื่องจาก X และ Y เป็นอิสระต่อกัน ทำให้ $f(x,y) = f_X(x)f_Y(y)$ ดังนั้น

$$E(XY)=\sum_x\sum_y xyf_X(x)f_Y(y)=\sum_x xf_X(x)\sum_y yf_Y(y)=E(X)E(Y)$$

กรณีที่ 2 ตัวแปรสุ่มแบบต่อเนื่อง

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy$$

เนื่องจาก X และ Y เป็นอิสระต่อกัน ทำให้ $f(x,y)=f_X(x)f_Y(y)$ ดังนั้น

$$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)dxdy=\int_{-\infty}^{\infty} xf_X(x)dx\int_{-\infty}^{\infty} yf_Y(y)dy=E(X)E(Y)$$

<u>**ตัวอย่างที่ 2.6.6**</u> กำหนดให้ X เป็นตัวแปรสุ่มแทนกระแสไฟฟ้าที่ไหลผ่านลวดทองแดงมีหน่วยเป็นมิลลิแอมแปร์ สมมติให้ X มีค่าอยู่ในช่วง [0, 20] และฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นดังนี้

$$f(x)=\begin{cases}0.05, & x\in[0,20]\\ 0, & x\notin[0,20]\end{cases}$$

จงหา

(1) ความน่าจะเป็นที่กระแสจะไหลผ่านลวดทองแดงน้อยกว่า 10 มิลลิแอมแปร์

(2) $E(X)$

(3) $E(X^2)$

(4) $E\left((3X-5)^2\right)$

(5) $E\left((X-E(X))^2\right)$

วิธีทำ พิจารณาตัวแปรสุ่ม X พบว่าเป็นตัวแปรสุ่มแบบต่อเนื่อง

(1) ความน่าจะเป็นที่กระแสจะไหลผ่านลวดทองแดงน้อยกว่า 10 มิลลิแอมแปร์ มีค่าเท่ากับ $P(X<10)$

$$P(X<10)=\int_{-\infty}^{10}0dx=\int_{-\infty}^{0}0dx+\int_{0}^{10}0.05dx=\left[0.05x\right]_{x=0}^{x=10}=0.5$$

ดังนั้นความน่าจะเป็นที่กระแสจะไหลผ่านลวดทองแดงน้อยกว่า 10 มิลลิแอมแปร์ มีค่าเท่ากับ 0.5

(2) จากนิยามค่าคาดหมายของตัวแปรสุ่มแบบต่อเนื่องหาได้จาก

$$\begin{aligned}E(X)&=\int_{-\infty}^{\infty}xf(x)dx\\&=\int_{-\infty}^{0}x\cdot 0dx+\int_{0}^{20}x\cdot 0.05dx+\int_{20}^{\infty}x\cdot 0dx\\&=0.05\left[\frac{x^2}{2}\right]_{x=0}^{20}=10\end{aligned}$$

ดังนั้นค่าคาดหมายของตัวแปรสุ่ม X มีค่าเท่ากับ 10

(3) จากนิยามค่าคาดหมายของฟังก์ชันที่เกี่ยวข้องกับตัวแปรสุ่มแบบต่อเนื่อง พิจารณา

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx \\ &= \int_{-\infty}^{0} x^2 \cdot 0dx + \int_{0}^{20} x^2 \cdot 0.05dx + \int_{20}^{\infty} x^2 \cdot 0dx \\ &= 0.05\left[\frac{x^3}{3}\right]_{x=0}^{20} = \frac{400}{3} \end{aligned}$$

ดังนั้นค่าคาดหมายของ X^2 มีค่าเท่ากับ $\frac{400}{3}$

(4) จากนิยามค่าคาดหมายของฟังก์ชันที่เกี่ยวข้องกับตัวแปรสุ่มแบบต่อเนื่อง และกฎของค่าคาดหมาย พิจารณา

$$\begin{aligned} E\left((3X-5)^2\right) &= E(9X^2 - 30X + 25) \\ &= 9E(X^2) - 30E(X) + 25 \\ &= 9\left(\frac{400}{3}\right) - 30(10) + 25 = 925 \end{aligned}$$

ดังนั้นค่าคาดหมายของ $(3X-5)^2$ มีค่าเท่ากับ 925

(5) จากนิยามค่าคาดหมายของฟังก์ชันที่เกี่ยวข้องกับตัวแปรสุ่มแบบต่อเนื่อง และกฎของค่าคาดหมาย พิจารณา

$$\begin{aligned} E\left((X-E(X))^2\right) &= E\left((X-10)^2\right) \\ &= E(X^2 - 20X + 100) \\ &= E(X^2) - 20E(X) + 100 \\ &= \frac{400}{3} - 20(10) + 100 = \frac{100}{3} \end{aligned}$$

ดังนั้นค่าคาดหมายของ $(X-E(X))^2$ มีค่าเท่ากับ $\frac{100}{3}$ ∎

2.7 ความแปรปรวนและส่วนเบี่ยงเบนมาตรฐาน

<u>นิยามที่ 2.7.1</u> *ความแปรปรวน (variance)* ของตัวแปรสุ่ม X คือ ค่าที่บอกถึงการกระจายของค่าตัวแปรสุ่ม X แทนด้วยสัญลักษณ์ $V(X)$ ซึ่งมีนิยามดังนี้

$$V(X) = E\left((X-E(X))^2\right)$$

พิจารณา

$$V(X)=E\left((X-E(X))^2\right)=E\left[X^2-2XE(X)+\left(E(X)\right)^2\right]$$

จากกฎของค่าคาดหมายจะได้ว่า

$$V(X)=E(X^2)-2E\left(XE(X)\right)+E\left(\left(E(X)\right)^2\right)$$

เนื่องจาก $E(X)$ และ $\left(E(X)\right)^2$ เป็นค่าคงที่ ใช้กฎของค่าคาดหมายจะได้ว่า

$$\begin{aligned}V(X)&=E(X^2)-2E(X)E(X)+\left(E(X)\right)^2\\&=E(X^2)-2\left(E(X)\right)^2+\left(E(X)\right)^2\\&=E(X^2)-\left(E(X)\right)^2\end{aligned}$$

ดังนั้นความแปรปรวนของตัวแปรสุ่ม X สามารถเขียนได้เป็น

$$V(X)=E(X^2)-\left(E(X)\right)^2$$

<u>หมายเหตุ</u>

1. ค่าความแปรปรวนมาก หมายความว่า ตัวแปรสุ่มนั้นมีการกระจายของข้อมูลมาก
2. ค่าความแปรปรวนน้อย หมายความว่า ตัวแปรสุ่มนั้นมีการกระจายของข้อมูลน้อย

กฎของความแปรปรวน

1. ความแปรปรวนของค่าคงที่มีค่าเท่ากับศูนย์ นั่นคือ

$$V(a)=0$$

สำหรับทุก $a\in\mathbb{R}$

<u>พิสูจน์</u> กำหนดให้ $a\in\mathbb{R}$ พิจารณา

$$V(a)=E\left((a-E(a))^2\right)=E\left((a-a)^2\right)=E(0)=0$$

2. ความแปรปรวนของตัวแปรสุ่มบวกกับค่าคงที่มีค่าเท่ากับความแปรปรวนของตัวแปรสุ่ม นั่นคือ

$$V(X+a)=V(X)$$

เมื่อ X เป็นตัวแปรสุ่มใด ๆ และ $a\in\mathbb{R}$

<u>พิสูจน์</u> กำหนดให้ X เป็นตัวแปรสุ่มใด ๆ และ $a\in\mathbb{R}$ พิจารณา

$$\begin{aligned}V(X+a)&=E\left((X+a-E(X+a))^2\right)\\&=E\left((X+a-E(X)-a)^2\right)\\&=E\left((X-E(X))^2\right)\\&=V(X)\end{aligned}$$

3. ความแปรปรวนของค่าคงที่คูณกับตัวแปรสุ่มมีค่าเท่ากับค่าคงที่ยกกำลังสองคูณกับความแปรปรวนของตัวแปรสุ่ม นั่นคือ

$$V(aX) = a^2V(X)$$

เมื่อ X เป็นตัวแปรสุ่มใด ๆ และ $a \in \mathbb{R}$

<u>พิสูจน์</u> กำหนดให้ X เป็นตัวแปรสุ่มใด ๆ และ $a \in \mathbb{R}$ พิจารณา

$$\begin{aligned} V(aX) &= E\left((aX - E(aX))^2\right) \\ &= E\left((aX - aE(X))^2\right) \\ &= E\left(a^2(X - E(X))^2\right) \\ &= a^2E\left((X - E(X))^2\right) \\ &= a^2V(X) \end{aligned}$$

4. ถ้า X และ Y เป็นตัวแปรสุ่มที่เป็นอิสระต่อกันแล้ว

$$V(aX \pm bY) = a^2V(X) + b^2V(Y)$$

เมื่อ $a, b \in \mathbb{R}$

<u>พิสูจน์</u> กำหนดให้ X และ Y เป็นตัวแปรสุ่มใด ๆ ที่เป็นอิสระต่อกัน และ $a, b \in \mathbb{R}$ พิจารณา

$$\begin{aligned} V(aX + bY) &= E\left((aX + bY - E(aX + bY))^2\right) \\ &= E\left((aX + bY - aE(X) - bE(Y))^2\right) \\ &= E\left(\left(a\left(X - E(X)\right) + b\left(Y - E(Y)\right)\right)^2\right) \\ &= E\left(a^2\left(X - E(X)\right)^2 + 2ab\left(X - E(X)\right)\left(Y - E(Y)\right) + b^2\left(Y - E(Y)\right)^2\right) \\ &= a^2E\left(\left(X - E(X)\right)^2\right) + 2abE\left(\left(X - E(X)\right)\left(Y - E(Y)\right)\right) + b^2E\left(\left(Y - E(Y)\right)^2\right) \\ &= a^2V(X) + 2abE\left(\left(X - E(X)\right)\left(Y - E(Y)\right)\right) + b^2V(Y) \\ &= a^2V(X) + 2abE\left(XY - XE(Y) - YE(X) + E(X)E(Y)\right) + b^2V(Y) \\ &= a^2V(X) + 2ab\left(E(XY) - E(X)E(Y)\right) + b^2V(Y) \end{aligned}$$

เนื่องจากตัวแปรสุ่ม X และ Y เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน ทำให้ $E(XY) = E(X)E(Y)$ ดังนั้น

$$V(aX + bY) = a^2V(X) + b^2V(Y)$$

ในทำนองเดียวกัน เราสามารถแสดงได้ว่า

$$V(aX - bY) = a^2V(X) + b^2V(Y)$$

<u>นิยามที่ 2.7.2</u> ***ส่วนเบี่ยงเบนมาตรฐาน (standard deviation)*** ของตัวแปรสุ่ม X แทนด้วยสัญลักษณ์ $SD(X)$ ซึ่งมีความสัมพันธ์กับความแปรปรวนตามความสมการต่อไปนี้

$$SD(X) = \sqrt{V(X)}$$

ตัวอย่างที่ 2.7.1 กำหนดให้ X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง ที่มีฟังก์ชันความน่าจะเป็นตามตารางต่อไปนี้

X	-5	-4	1	2
$f(x)$	1/4	1/8	1/2	1/8

และ $f(x)=0$ เมื่อ $x \neq -5,-4,1,2$ จงหา $V(X)$ และ $SD(X)$

วิธีทำ เนื่องจากความแปรปรวนของตัวแปรสุ่ม X หาได้จากสมการ

$$V(X)=E(X^2)-[E(X)]^2$$

พิจารณาการหาค่า $E(X)$ ของตัวแปรสุ่มแบบไม่ต่อเนื่อง ได้จากสมการ

$$E(X)=\sum_x xf(x)=(-5)f(-5)+(-4)f(-4)+(1)f(1)+(2)f(2)$$

$$=(-5)\left(\frac{1}{4}\right)+(-4)\left(\frac{1}{8}\right)+\frac{1}{2}+2\left(\frac{1}{8}\right)=-1$$

จากนั้นหาค่า $E(X^2)$ ได้ดังนี้

$$E(X^2)=\sum_x x^2 f(x)=(-5)^2 f(-5)+(-4)^2 f(-4)+1^2 f(1)+2^2 f(2)$$

$$=25\left(\frac{1}{4}\right)+16\left(\frac{1}{8}\right)+\frac{1}{2}+4\left(\frac{1}{8}\right)=\frac{37}{4}$$

แทนค่า $E(X^2)$ และ $E(X)$ เพื่อหาค่า $V(X)$ จะได้ว่า

$$V(X)=E(X^2)-[E(X)]^2=\frac{37}{4}-(-1)^2=8.25$$

ส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X หาได้จาก

$$SD(X)=\sqrt{V(X)}=\sqrt{8.25}\approx 2.87$$

ดังนั้นความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X มีค่าเท่ากับ 8.25 และ 2.87 ตามลำดับ ∎

ตัวอย่างที่ 2.7.2 กำหนดให้ X เป็นตัวแปรสุ่มแบบต่อเนื่อง ที่มีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x)=\begin{cases}\frac{x}{2}, & x\in(0,2)\\ 0, & x\notin(0,2)\end{cases}$$

จงหาความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X

วิธีทำ เนื่องจากความแปรปรวนของตัวแปรสุ่ม X หาได้จากสมการ

$$V(X)=E(X^2)-[E(X)]^2$$

พิจารณาการหาค่า $E(X)$ ของตัวแปรสุ่มแบบต่อเนื่อง ได้จากสมการ

$$E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_{-\infty}^{0}x\cdot 0dx+\int_{0}^{2}x\cdot\frac{x}{2}dx+\int_{2}^{\infty}x\cdot 0dx=\int_{0}^{2}\frac{x^2}{2}dx=\frac{4}{3}$$

จากนั้นหาค่า $E(X^2)$ ได้ดังนี้

$$E(X^2)=\int_{-\infty}^{\infty}x^2f(x)dx=\int_{-\infty}^{0}x^2\cdot 0dx+\int_{0}^{2}x^2\cdot\frac{x}{2}dx+\int_{2}^{\infty}x^2\cdot 0dx=\int_{0}^{2}\frac{x^3}{2}dx=2$$

แทนค่า $E(X^2)$ และ $E(X)$ เพื่อหาค่า $V(X)$ จะได้ว่า

$$V(X)=E(X^2)-[E(X)]^2=2-\left(\frac{4}{3}\right)^2=\frac{2}{9}$$

ส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X หาได้จาก

$$SD(X)=\sqrt{V(X)}=\sqrt{2/9}\approx 0.47$$

ดังนั้นความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X มีค่าเท่ากับ $\frac{2}{9}$ และ 0.47 ตามลำดับ ■

<u>ตัวอย่างที่ 2.7.3</u> พนักงานขายสินค้าคนหนึ่งทำหน้าที่ขายสินค้า 2 ชนิดคือ A และ B โดยที่ยอดขายสินค้าทั้งสองชนิดเป็นอิสระต่อกัน เงินเดือนของพนักงานได้จากยอดขาย 20% ของสินค้า A และ 30% ของสินค้า B จากข้อมูลในอดีตพบว่าเขาขายสินค้า A ได้โดยเฉลี่ยเดือนละ 100,000 บาท ส่วนเบี่ยงเบนมาตรฐาน 8,000 บาท เขาขายสินค้า B ได้เฉลี่ยเดือนละ 60,000 บาท ส่วนเบี่ยงเบนมาตรฐาน 3,000 บาท จงหาเงินเดือนที่คาดว่าจะได้และส่วนเบี่ยงเบนมาตรฐานของเงินเดือนของพนักงานคนนี้

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนเงินเดือนของพนักงานคนนี้

A แทนยอดขายสินค้าชนิด A

B แทนยอดขายสินค้าชนิด B

สามารถเขียนความสัมพันธ์ระหว่าง A, B และตัวแปรสุ่ม X ได้ดังนี้

$$X=\frac{20}{100}A+\frac{30}{100}B$$

ดังนั้นเงินเดือนที่คาดว่าจะได้รับของพนักงานคนนี้มีค่าเท่ากับ $E(X)$

$$E(X)=E\left(\frac{20}{100}A+\frac{30}{100}B\right)=\frac{20}{100}E(A)+\frac{30}{100}E(B)$$

$$=\frac{20}{100}(100,000)+\frac{30}{100}(60,000)=38,000$$

และส่วนเบี่ยงเบนมาตรฐานของเงินเดือนของพนักงานคนนี้มีค่าเท่ากับ $SD(X)$

$$V(X)=V\left(\frac{20}{100}A+\frac{30}{100}B\right)=\left(\frac{20}{100}\right)^2V(A)+\left(\frac{30}{100}\right)^2V(B)$$

$$=\left(\frac{20}{100}\right)^2(8,000)^2+\left(\frac{30}{100}\right)^2(3,000)^2=3,370,000$$

นั่นคือ

$$SD(X) = \sqrt{3,370,000} \approx 1,835.76$$

ดังนั้นเงินเดือนที่คาดว่าจะได้และส่วนเบี่ยงเบนมาตรฐานของเงินเดือนพนักงานคนนี้มีค่าเท่ากับ 38,000 บาท และ 1,835.76 บาท ตามลำดับ ■

บทที่ 3
การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่อง (Discrete Probability Distributions)

ในบทที่ 3 นี้จะกล่าวถึงการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่องที่สำคัญ ได้แก่ การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มไม่ต่อเนื่อง การแจกแจงความน่าจะเป็นแบบเบอร์นูลี การแจกแจงความน่าจะเป็นแบบทวินาม และการแจกแจงความน่าจะเป็นแบบปัวซง

3.1 การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มไม่ต่อเนื่อง

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่องที่ง่ายและซับซ้อนน้อยที่สุด คือ การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มไม่ต่อเนื่อง ซึ่งจะได้กล่าวถึงในหัวข้อนี้

นิยามที่ 3.1.1 ถ้าตัวแปรสุ่ม X เป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง มีสมาชิกในเรนจ์ทั้งหมด คือ $x_1,\dots,x_n$ และแต่ละค่า x_i มีความน่าจะเป็นเท่า ๆ กันแล้ว จะเรียก X ว่าเป็นตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มไม่ต่อเนื่อง (discrete uniform probability distribution)*** ดังนั้น ถ้า X เป็นตัวแปรสุ่มแบบยูนิฟอร์มไม่ต่อเนื่องแล้ว จะเขียนแทนได้ด้วยสัญลักษณ์

$$X \sim U(n)$$

โดยที่ n แทนจำนวนสมาชิกในเรนจ์ทั้งหมดของตัวแปรสุ่ม X (หรือจำนวนค่าที่เป็นไปได้ทั้งหมดของตัวแปรสุ่ม X)

พิจารณา $X \sim U(n)$ จากนิยามที่ 3.1.1 จะได้ว่า

$$P(X = x_i) = \frac{1}{n}$$

สำหรับทุกค่า $i = 1,2,\dots,n$

$$X \sim U(6)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการนี้

$$f(x,6) = \frac{1}{6}$$ ■

ตัวอย่างที่ 3.1.2 ในกระบวนการเคลือบผิววัตถุชนิดหนึ่ง ความหนาของการเคลือบผิวมีค่าที่เป็นไปได้คือ 0.15, 0.16, 0.17, 0.18 และ 0.19 ด้วยความน่าจะเป็นเท่า ๆ กัน จงหา

(1) ค่าเฉลี่ยและความแปรปรวนความหนาของการเคลือบผิวนี้

(2) ความน่าจะเป็นที่ความหนาของการเคลือบผิวมีค่ามากกว่า 0.16

(3) ความน่าจะเป็นที่ความหนาของการเคลือบผิวมีค่าไม่เกิน 0.16

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนความหนาของการเคลือบผิว เนื่องจากค่าที่เป็นไปได้ของตัวแปรสุ่ม X มีทั้งหมด 5 ค่า และความน่าจะเป็นของแต่ละค่ามีค่าเท่ากัน ดังนั้น X จึงเป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มไม่ต่อเนื่อง แทนได้ด้วยสัญลักษณ์

$$X \sim U(5)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการนี้

$$f(x,5) = \frac{1}{5}$$

(1) ค่าเฉลี่ยความหนาของการเคลือบผิวนี้หาได้จาก

$$\begin{aligned} E(X) &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ &= \frac{1}{5}\sum_{i=1}^{5} x_i \\ &= \frac{1}{5}(0.15+0.16+0.17+0.18+0.19) = 0.17 \end{aligned}$$

ความแปรปรวนความหนาของการเคลือบผิวนี้หาได้จาก

$$\begin{aligned} V(X) &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \\ &= \frac{1}{5}\sum_{i=1}^{5}(x_i-0.17)^2 \\ &= \frac{1}{5}\left((0.15-0.17)^2+(0.16-0.17)^2+(0.17-0.17)^2+(0.18-0.17)^2+(0.19-0.17)^2\right) \\ &= 0.0002 \end{aligned}$$

ดังนั้น ค่าเฉลี่ยและความแปรปรวนความหนาของการเคลือบผิวนี้มีค่าเท่ากับ 0.17 และ 0.0002 ตามลำดับ

(2) ความน่าจะเป็นที่ความหนาของการเคลือบผิวมีค่ามากกว่า 0.16 แทนได้ด้วยสัญลักษณ์ $P(X>0.16)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X>0.16) &= P(X=0.17)+P(X=0.18)+P(X=0.19) \\ &= f(0.17)+f(0.18)+f(0.19) \\ &= \frac{1}{5}+\frac{1}{5}+\frac{1}{5}=\frac{3}{5}=0.6 \end{aligned}$$

ดังนั้น ความหนาของการเคลือบผิวมีค่ามากกว่า 0.16 ด้วยความน่าจะเป็นเท่ากับ 0.6

(3) ความน่าจะเป็นที่ความหนาของการเคลือบผิวมีค่าไม่เกิน 0.16 มีค่าเท่ากับ $P(X \le 0.16)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X \le 0.16) &= P(X=0.15)+P(X=0.16) \\ &= f(0.15)+f(0.16) \\ &= \frac{1}{5}+\frac{1}{5}=\frac{2}{5}=0.4 \end{aligned}$$

ดังนั้น ความหนาของการเคลือบผิวมีค่าไม่เกิน 0.16 ด้วยความน่าจะเป็นเท่ากับ 0.4 ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 3.1.2 ข้อที่ (3) เราสามารถหาค่า $P(X \le 0.16)$ ได้อีกทางหนึ่งโดยใช้คำตอบของข้อที่ (2) ดังนี้

$$P(X \le 0.16) = 1-P(X>0.16) = 1-0.6 = 0.4$$

3.2 การแจกแจงความน่าจะเป็นแบบเบอร์นูลี

ในหัวข้อนี้จะกล่าวถึง การแจกแจงความน่าจะเป็นแบบเบอร์นูลี ซึ่งเป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่องที่สำคัญและเป็นพื้นฐานสำหรับการแจกแจงความน่าจะเป็นแบบอื่นอีกด้วย

<u>นิยามที่ 3.2.1</u> การทดลองที่มีผลลัพธ์เพียง 2 แบบ คือ สิ่งที่สนใจ (success) และสิ่งที่ไม่สนใจ (failure) โดยที่ความน่าจะเป็นในการเกิดสิ่งที่สนใจ และสิ่งที่ไม่สนใจมีค่าคงที่เสมอ เรียกการทดลองดังกล่าวว่า ***การทดลองแบบเบอร์นูลี (Bernoulli experiment)***

<u>นิยามที่ 3.2.2</u> ถ้า X เป็นตัวแปรสุ่มที่สอดคล้องกับการทดลองแบบเบอร์นูลี จะเรียก X ว่าเป็นตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบเบอร์นูลี (Bernoulli probability distribution)*** เขียนแทนได้ด้วยสัญลักษณ์

$$X \sim Ber(p)$$

เมื่อ p แทนความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ

พิจารณา $X \sim Ber(p)$ จากนิยามที่ 3.2.1 และ 3.2.2 เราสามารถนิยามตัวแปรสุ่ม X ได้ดังนี้

$$X = \begin{cases} 1, & S \\ 0, & F \end{cases}$$

โดยที่ S แทนการเกิดเหตุการณ์ที่สนใจซึ่งมีความน่าจะเป็นในการเกิดเท่ากับ p

F แทนการเกิดเหตุการณ์ที่ไม่สนใจซึ่งมีความน่าจะเป็นในการเกิดเท่ากับ $1-p$

ข้อสังเกต

ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี จะมีค่าที่เป็นไปได้แค่ 2 ค่าเท่านั้น และความน่าจะเป็นของทั้ง 2 ค่ามีค่าคงที่เสมอ

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี

พิจารณา $X \sim Ber(p)$ จากนิยามของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี ดังนั้น ฟังก์ชันความน่าจะเป็น $f(x,p)$ หาได้จากสมการ

$$f(x,p) = P(X = x) = \begin{cases} 1-p, & x = 0 \\ p, & x = 1 \end{cases}$$

หรือ

$$f(x,p) = p^x (1-p)^{1-x}$$

โดยที่ $x = 0,1$ และ $0 \le p \le 1$

หมายเหตุ

ในหัวข้อนี้ใช้สัญลักษณ์ $f(x,p)$ แทน $f(x)$ เนื่องจากจะแสดงให้เห็นว่าฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลีมีความสัมพันธ์กับความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ

ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี

พิจารณา $X \sim Ber(p)$ จากนิยามค่าคาดหมายหรือค่าเฉลี่ยของตัวแปรสุ่มแบบไม่ต่อเนื่องจะได้ว่า

$$\begin{aligned} E(X) &= \sum_x x f(x,p) \\ &= 0 f(0,p) + 1 f(1,p) \\ &= 0(1-p) + 1(p) = p \end{aligned}$$

ดังนั้น ค่าเฉลี่ยหรือค่าคาดหมายของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลีมีค่าเท่ากับ

$$E(X) = p$$

ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี

พิจารณา $X \sim Ber(p)$ จากนิยามความแปรปรวนของตัวแปรสุ่มจะได้ว่า

$$\begin{aligned} V(X) &= E[(X-E(X))^2] \\ &= E[(X-p)^2] \\ &= \sum_x (x-p)^2 f(x,p) \\ &= (0-p)^2 f(0,p) + (1-p)^2 f(1,p) \\ &= (0-p)^2(1-p) + (1-p)^2(p) \\ &= p^2(1-p) + p(1-p)^2 \\ &= p(1-p)(p+(1-p)) \\ &= p(1-p) \end{aligned}$$

ดังนั้น ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงแบบความน่าจะเป็นเบอร์นูลีมีค่าเท่ากับ

$$V(X) = p(1-p)$$

ตัวอย่างที่ 3.2.1

(1) พิจารณาการทดลองสุ่มโยนเหรียญหนึ่งเหรียญหนึ่งครั้ง โดยที่ความน่าจะเป็นในการที่เหรียญ ๆ นี้จะออกหัวมีค่าเท่ากับ 0.4 กำหนดให้ X แทนจำนวนครั้งที่เหรียญออกหัว

เนื่องจากการโยนเหรียญเกิดผลลัพธ์ได้เพียง 2 แบบ คือ ออกหัว หรือก้อยเท่านั้น โดยที่สิ่งที่เราสนใจคือการออกหัว และความน่าจะเป็นในการออกหัวหรือก้อยก็มีค่าคงที่เสมอ นั่นคือ การโยนเหรียญหนึ่งเหรียญหนึ่งครั้งจึงเป็นการทดลองแบบเบอร์นูลี ดังนั้น ตัวแปรสุ่ม X จึงเป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี แทนได้ด้วยสัญลักษณ์

$$X \sim Ber(0.4)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการนี้

$$f(x,0.4) = \begin{cases} 0.6, & x=0 \\ 0.4, & x=1 \end{cases}$$

หรือ

$$f(x,0.4) = 0.4^x(1-0.4)^{1-x}$$

โดยที่ $x = 0,1$

(2) พิจารณาการทดลองสุ่มหยิบลูกบอล 1 ลูก จากถุงใบหนึ่งซึ่งมีลูกบอลสีแดง M ลูก สีขาว N ลูก กำหนดให้ X แทนจำนวนลูกบอลสีขาวที่หยิบได้

เนื่องจากการหยิบลูกบอลเกิดผลลัพธ์ได้เพียง 2 แบบ คือ หยิบได้ลูกบอลสีขาว หรือสีแดงเท่านั้น โดยที่สิ่งที่เราสนใจคือการหยิบได้ลูกบอลสีขาว และความน่าจะเป็นในการหยิบได้ลูกบอลสีขาวหรือสีแดงก็มีค่าคงที่เสมอ สามารถหาค่าความน่าจะเป็นได้ดังนี้

$$P(\text{การหยิบได้ลูกบอลสีขาว}) = \frac{N}{N+M}$$

$$P(\text{การหยิบได้ลูกบอลสีแดง}) = \frac{M}{N+M}$$

นั่นคือ การหยิบลูกบอล 1 ลูก จากถุงใบหนึ่งซึ่งมีลูกบอลสีแดง M ลูก สีขาว N ลูก จึงเป็นการทดลองแบบเบอร์นูลี ดังนั้น ตัวแปรสุ่ม X จึงเป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี แทนได้ด้วยสัญลักษณ์

$$X \sim Ber\left(\frac{N}{N+M}\right)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการนี้

$$f(x,\frac{N}{N+M}) = \begin{cases} \dfrac{M}{N+M}, & x=0 \\ \dfrac{N}{N+M}, & x=1 \end{cases}$$

หรือ

$$f(x,\frac{N}{N+M}) == \left(\frac{N}{N+M}\right)^{x}\left(1-\frac{N}{N+M}\right)^{1-x}$$

โดยที่ $x = 0,1$

(3) พิจารณาการทดลองสุ่มหยิบสินค้ามาตรวจสอบ 1 ชิ้น จากการผลิตสินค้าทั้งหมด N ชิ้น ที่มีสินค้าที่ไม่ได้มาตรฐาน M ชิ้น กำหนดให้ X แทนจำนวนสินค้าที่ไม่ได้มาตรฐาน

เนื่องจากการหยิบสินค้าเกิดผลลัพธ์ได้เพียง 2 แบบ คือ หยิบได้สินค้าที่ได้มาตรฐาน หรือไม่ได้มาตรฐานเท่านั้น โดยที่สิ่งที่เราสนใจคือการหยิบได้สินค้าที่ไม่ได้มาตรฐาน และความน่าจะเป็นในการหยิบได้สินค้าที่ไม่ได้มาตรฐานหรือสินค้าที่ได้มาตรฐานก็มีค่าคงที่เสมอ สามารถหาค่าความน่าจะเป็นได้ดังนี้

$$P\text{(การหยิบได้สินค้าที่ไม่ได้มาตรฐาน)} = \frac{M}{N}$$

$$P\text{(การหยิบได้สินค้าที่ได้มาตรฐาน)} = \frac{N-M}{N}$$

นั่นคือ การหยิบสินค้า 1 ชิ้น จากการผลิตสินค้าทั้งหมด N ชิ้น ที่มีสินค้าที่ไม่ได้มาตรฐาน M ชิ้น จึงเป็นการทดลองแบบเบอร์นูลี ดังนั้น ตัวแปรสุ่ม X จึงเป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี แทนได้ด้วยสัญลักษณ์

$$X \sim Ber\left(\frac{M}{N}\right)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการนี้

$$f(x,\frac{M}{N}) = \begin{cases} \dfrac{N-M}{N}, & x=0 \\ \dfrac{M}{N}, & x=1 \end{cases}$$

หรือ

$$f(x,\frac{M}{N}) = \left(\frac{M}{N}\right)^{x}\left(1-\frac{M}{N}\right)^{1-x}$$

โดยที่ $x = 0,1$ ■

<u>ตัวอย่างที่ 3.2.2</u> สุ่มหยิบสินค้า 1 ชิ้น จากกล่องที่มีสินค้าที่แตกต่างกัน 10 ชิ้น ซึ่งมีสินค้าที่ชำรุดปนอยู่ 4 ชิ้น จงหา

(1) ค่าเฉลี่ยและความแปรปรวนของการหยิบได้สินค้าที่ชำรุด

(2) ค่าเฉลี่ยและความแปรปรวนของการหยิบได้สินค้าที่ไม่ชำรุด

วิธีทำ เนื่องจากการหยิบสินค้าจากกล่องเกิดผลลัพธ์ได้เพียง 2 แบบ คือ หยิบได้สินค้าที่ชำรุด หรือหยิบได้สินค้าที่ไม่ชำรุดเท่านั้น และความน่าจะเป็นในการหยิบได้สินค้าที่ชำรุดหรือหยิบได้สินค้าที่ไม่ชำรุดมีค่าคงที่เสมอ สามารถหาค่าความน่าจะเป็นได้ดังนี้

$$P(\text{การหยิบได้สินค้าที่ชำรุด}) = \frac{4}{10} = 0.4$$

$$P(\text{การหยิบได้สินค้าที่ไม่ชำรุด}) = \frac{6}{10} = 0.6$$

นั่นคือ การหยิบสินค้า 1 ชิ้น จากกล่องที่มีสินค้าที่แตกต่างกัน 10 ชิ้น ซึ่งมีสินค้าที่ชำรุดปนอยู่ 4 ชิ้น จึงเป็นการทดลองแบบเบอร์นูลี

(1) กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนสินค้าที่ชำรุด ดังนั้น ตัวแปรสุ่ม X จึงเป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี แทนได้ด้วยสัญลักษณ์

$$X \sim Ber(0.4)$$

ค่าเฉลี่ยของการหยิบได้สินค้าที่ชำรุดมีค่าเท่ากับ

$$E(X) = p = 0.4$$

ความแปรปรวนของการหยิบได้สินค้าที่ชำรุดมีค่าเท่ากับ

$$V(X) = p(1-p) = 0.4(1-0.4) = (0.4)(0.6) = 0.24$$

ดังนั้น ค่าเฉลี่ยและความแปรปรวนของการหยิบได้สินค้าที่ชำรุดมีค่าเท่ากับ 0.4 และ 0.24 ตามลำดับ

(2) กำหนดให้ Y เป็นตัวแปรสุ่มแทนจำนวนสินค้าที่ไม่ชำรุด ดังนั้น ตัวแปรสุ่ม Y จึงเป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลี แทนได้ด้วยสัญลักษณ์

$$Y \sim Ber(0.6)$$

ค่าเฉลี่ยของการหยิบได้สินค้าที่ไม่ชำรุดมีค่าเท่ากับ

$$E(X) = p = 0.6$$

ความแปรปรวนของการหยิบได้สินค้าที่ไม่ชำรุดมีค่าเท่ากับ

$$V(X) = p(1-p) = 0.6(1-0.6) = (0.6)(0.4) = 0.24$$

ดังนั้น ค่าเฉลี่ยและความแปรปรวนของการหยิบได้สินค้าที่ไม่ชำรุดจากกล่องใบนี้มีค่าเท่ากับ 0.6 และ 0.24 ตามลำดับ ■

หมายเหตุ

การทดลองแบบไม่ใส่คืนไม่ถือว่าเป็นการทดลองแบบทวินาม เนื่องจากการทดลองแต่ละครั้งมีผลกระทบต่อการทดลองครั้งต่อ ๆ ไป (นั่นคือ การทดลองแต่ละครั้งไม่เป็นอิสระต่อกัน) แต่ถ้าจำนวนสิ่งของที่สุ่มมามีจำนวนน้อยมากเมื่อเทียบกับสิ่งของทั้งหมด ก็สามารถอนุโลมว่าการทดลองแบบไม่ใส่คืนเป็นการทดลองแบบทวินามได้

ตัวอย่างที่ 3.3.2 กล่องใหญ่ใบหนึ่งบรรจุสินค้าจำนวนมาก โดยที่ในกล่องใบนี้มีสินค้าชำรุดปนอยู่ 5% ถ้าสุ่มตัวอย่างสินค้าจากกล่องนี้มา 10 ชิ้น โดยสุ่มมาทีละชิ้นแล้วใส่คืน จงหา

(1) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดอย่างน้อย 4 ชิ้น
(2) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดมากกว่า 1 ชิ้น
(3) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดตั้งแต่ 7 ชิ้น ขึ้นไป
(4) จำนวนสินค้าชำรุดที่คาดไว้
(5) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ไม่ชำรุดมากกว่า 3 ชิ้น

วิธีทำ พิจารณาการทดลองสุ่มเลือกสินค้ามา 1 ชิ้น ผลลัพธ์ที่ได้มีเพียง 2 แบบ เท่านั้น คือ สุ่มได้สินค้าที่ชำรุด หรือสุ่มได้สินค้าที่ไม่ชำรุด สิ่งที่สนใจคือการสุ่มได้สินค้าที่ชำรุด ซึ่งมีความน่าจะเป็นคงที่เสมอ นั่นคือการสุ่มเลือกสินค้า 1 ชิ้น เป็นการทดลองแบบเบอร์นูลี ดังนั้นการสุ่มเลือกสินค้ามา 10 ชิ้น ถือว่าเป็นการทดลองแบบเบอร์นูลีซ้ำกัน 10 ครั้ง การทดลองแต่ละครั้งเป็นอิสระต่อกัน ดังนั้นจึงถือว่าเป็นการทดลองแบบทวินาม

กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนสินค้าที่ชำรุด จากโจทย์จะได้ว่า $n=10$ และ p เท่ากับความน่าจะเป็นที่จะพบสินค้าที่ชำรุด นั่นคือ $p=0.05$ ดังนั้น

$$X \sim B(10, 0.05)$$

(1) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดอย่างน้อย 4 ชิ้น แทนได้ด้วยสัญลักษณ์ $P(X \geq 4)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X \geq 4) &= P(X=4) + \ldots + P(X=10) \\ &= \binom{10}{4}(0.05)^4(1-0.05)^{10-4} + \ldots + \binom{10}{10}(0.05)^{10}(1-0.05)^{10-10} \\ &= 0.0010 + 0.0001 + 0 + 0 + 0 + 0 + 0 \\ &= 0.0011 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดอย่างน้อย 4 ชิ้น มีค่าเท่ากับ 0.0011

(2) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดมากกว่า 1 ชิ้น แทนได้ด้วยสัญลักษณ์ $P(X > 1)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X > 1) &= P(X=2) + P(X=3) + \ldots + P(X=10) \\ &= \binom{10}{2}(0.05)^2(1-0.05)^{10-2} + \ldots + \binom{10}{10}(0.05)^{10}(1-0.05)^{10-10} \\ &= 0.0746 + 0.0105 + 0.0010 + 0.0001 + 0 + 0 + 0 + 0 + 0 \end{aligned}$$

$$= 0.0862$$

ดังนั้น ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดมากกว่า 1 ชิ้น มีค่าเท่ากับ 0.0862

(3) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดตั้งแต่ 7 ชิ้น ขึ้นไป แทนได้ด้วยสัญลักษณ์ $P(X \geq 7)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X \geq 7) &= P(X = 7) + \ldots + P(X = 10) \\ &= \binom{10}{7}(0.05)^7(1-0.05)^{10-7} + \ldots + \binom{10}{10}(0.05)^{10}(1-0.05)^{10-10} \\ &= 0 + 0 + 0 + 0 = 0 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ชำรุดตั้งแต่ 7 ชิ้น ขึ้นไป มีค่าเท่ากับ 0

(4) จำนวนสินค้าชำรุดที่คาดไว้หาได้จาก

$$E(X) = np = 10(0.05) = 0.5$$

ดังนั้น จำนวนสินค้าชำรุดที่คาดไว้มีค่าเท่ากับ 0.5 ชิ้น

กำหนดให้ Y เป็นตัวแปรสุ่มแทนจำนวนสินค้าที่ไม่ชำรุด ทำให้สิ่งที่สนใจเปลี่ยนเป็นสินค้าที่ไม่ชำรุดแทน จากโจทย์จะได้ว่า $n = 10$ และ p เท่ากับความน่าจะเป็นที่จะพบสินค้าที่ไม่ชำรุด นั่นคือ $p = 0.95$ ดังนั้น

$$Y \sim B(10, 0.95)$$

(5) ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ไม่ชำรุดมากกว่า 3 ชิ้น แทนได้ด้วยสัญลักษณ์ $P(Y > 3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(Y > 3) &= P(Y = 4) + \ldots + P(Y = 10) \\ &= \binom{10}{4}(0.95)^4(1-0.95)^{10-4} + \ldots + \binom{10}{10}(0.95)^{10}(1-0.95)^{10-10} \\ &= 0 + 0.0010 + 0.0001 + 0.0105 + 0.0746 + 0.3151 + 0.5987 = 1 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่จะสุ่มได้สินค้าที่ไม่ชำรุดมากกว่า 3 ชิ้น มีค่าเท่ากับ 1 ■

<u>หมายเหตุ</u>

การคำนวณหาค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทวินาม สามารถหาค่าได้โดยตรงจากฟังก์ชันความน่าจะเป็น หรือใช้ตารางที่ 1 ในภาคผนวก

<u>ตัวอย่างที่ 3.3.3</u> กล่องใหญ่ใบหนึ่งบรรจุสินค้าจำนวนมาก โดยที่ในกล่องใบนี้มีสินค้าชำรุดปนอยู่ 5% ถ้าสุ่มตัวอย่างสินค้าจากกล่องนี้มา 10 ชิ้น พบว่ามีสินค้าที่ชำรุดตั้งแต่สองชิ้นขึ้นไป จะไม่ยอมรับสินค้านั้นทั้งกล่อง จงหา

(1) ความน่าจะเป็นที่จะยอมรับสินค้าทั้งกล่อง

(2) ความน่าจะเป็นที่จะไม่ยอมรับสินค้าทั้งกล่อง

วิธีทำ ตัวอย่างนี้จะแตกต่างจากตัวอย่างที่ 3.3.2 เล็กน้อยตรงที่เป็นการสุ่มแบบไม่ใส่คืน แต่เนื่องจากโจทย์กำหนดว่าสินค้าทั้งหมดมีจำนวนมาก ดังนั้นจึงถือว่าเป็นการทดลองแบบทวินาม

กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนสินค้าที่ชำรุด จากโจทย์จะได้ว่า $n=10$ และ p เท่ากับความน่าจะเป็นที่จะพบสินค้าที่ชำรุด นั่นคือ $p=0.05$ ดังนั้น

$$X \sim B(10, 0.05)$$

ดังนั้น การหาค่าความน่าจะเป็นของเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม X สามารถหาได้โดยใช้ตารางที่ 1.10 ในภาคผนวก

(1) การที่จะยอมรับสินค้าทั้งกล่องมีเงื่อนไขว่า จากการสุ่มสินค้ามา 10 ชิ้น จะต้องพบสินค้าที่ชำรุดไม่ถึง 2 ชิ้น ดังนั้น ความน่าจะเป็นที่จะยอมรับสินค้าทั้งกล่องหาค่าได้ดังนี้

$$\begin{aligned} P(\text{การยอมรับสินค้าทั้งกล่อง}) &= P(X<2) \\ &= P(X=0)+P(X=1) \\ &= 0.5987+0.3151 \\ &= 0.9138 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่ยอมรับสินค้าทั้งกล่องมีค่าเท่ากับ 0.9138

(2) การที่จะไม่ยอมรับสินค้าทั้งกล่องมีเงื่อนไขว่า จากการสุ่มสินค้ามา 10 ชิ้น จะต้องพบสินค้าที่ชำรุดตั้งแต่ 2 ชิ้นขึ้นไป ดังนั้น ความน่าจะเป็นที่จะไม่ยอมรับสินค้าทั้งกล่องหาค่าได้ดังนี้

$$\begin{aligned} P(\text{การไม่ยอมรับสินค้าทั้งกล่อง}) &= P(X\geq 2) \\ &= P(X=2)+\ldots+P(X=10) \\ &= 1-\{P(X=0)+P(X=1)\} \\ &= 1-0.9138 \\ &= 0.0862 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่จะไม่ยอมรับสินค้าทั้งกล่องมีค่าเท่ากับ 0.0862 ■

<u>ข้อสังเกต</u>

จากตัวอย่างที่ 3.3.3 ข้อที่ (2) $P(\text{การไม่ยอมรับสินค้าทั้งกล่อง}) = 1-P(\text{การยอมรับสินค้าทั้งกล่อง})$ ดังนั้นเราสามารถหาคำตอบได้เลยเมื่อทราบคำตอบในข้อที่ (1)

<u>ตัวอย่างที่ 3.3.4</u> ความน่าจะเป็นที่คนไข้จะหายจากการเป็นโรคชนิดหนึ่งหลังจากได้รับการรักษาแล้วมีค่าเท่ากับ 0.4 ถ้าสุ่มคนไข้ที่เป็นโรคชนิดนี้มา 10 คน จากคนไข้จำนวนมาก จงหา

(1) ความน่าจะเป็นที่มีผู้หายจากโรคนี้อย่างน้อยที่สุด 9 คน

(2) ความน่าจะเป็นที่มีผู้หายจากโรคนี้ไม่เกิน 9 คน

(3) ความน่าจะเป็นที่มีผู้หายจากโรคนี้ไม่ถึง 9 คน

(4) ความน่าจะเป็นที่มีผู้ไม่หายจากโรคนี้น้อยกว่า 1 คน

(5) ความน่าจะเป็นที่มีผู้ไม่หายจากโรคนี้ตั้งแต่ 8 คนขึ้นไป

วิธีทำ พิจารณาการทดลองสุ่มคนไข้มา 1 คน ผลลัพธ์ที่ได้มีเพียง 2 แบบ เท่านั้น คือ สุ่มได้คนไข้ที่หายจากโรค หรือสุ่มได้คนไข้ที่ไม่หายจากโรค สิ่งที่สนใจคือการสุ่มได้คนไข้ที่หายจากโรค ซึ่งมีความน่าจะเป็นคงที่เสมอ นั่นคือ การสุ่มคนไข้มา 1 คน เป็นการทดลองแบบเบอร์นูลี ดังนั้นการสุ่มเลือกคนไข้มา 10 คน ถือว่าเป็นการทดลองแบบเบอร์นูลีซ้ำกัน 10 ครั้ง ถ้าเป็นการสุ่มแบบใส่คืน การทดลองแต่ละครั้งเป็นอิสระต่อกัน จะถือว่าเป็นการทดลองแบบทวินาม สำหรับโจทย์ข้อนี้เป็นการสุ่มแบบไม่ใส่คืน แต่เนื่องจากสุ่มจากคนไข้จำนวนมาก ดังนั้นจึงถือว่าเป็นการทดลองแบบทวินาม

กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนคนไข้ที่หายจากโรคนี้หลังได้รับการรักษา จากโจทย์จะได้ว่า $n=10$ และ p เท่ากับความน่าจะเป็นที่คนไข้จะหายจากการเป็นโรคนี้หลังได้รับการรักษา นั่นคือ $p=0.4$ ดังนั้น

$$X \sim B(10,0.4)$$

ดังนั้น การหาค่าความน่าจะเป็นของเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม X สามารถหาได้โดยใช้ตารางที่ 1.10 ในภาคผนวก

(1) ความน่าจะเป็นที่มีผู้หายจากโรคนี้อย่างน้อยที่สุด 9 คน แทนได้ด้วยสัญลักษณ์ $P(X \geq 9)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X \geq 9) &= P(X=9)+P(X=10) \\ &= 0.0016+0.0001 \\ &= 0.0017 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่มีผู้หายจากโรคนี้อย่างน้อยที่สุด 9 คน มีค่าเท่ากับ 0.0017

(2) ความน่าจะเป็นที่มีผู้หายจากโรคนี้ไม่เกิน 9 คน แทนได้ด้วยสัญลักษณ์ $P(X \leq 9)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X \leq 9) &= P(X=0)+\ldots+P(X=9) \\ &= 1-P(X=10) \\ &= 1-0.0001 \\ &= 0.9999 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่มีผู้หายจากโรคนี้ไม่เกิน 9 คน มีค่าเท่ากับ 0.9999

(3) ความน่าจะเป็นที่มีผู้หายจากโรคนี้ไม่ถึง 9 คน แทนได้ด้วยสัญลักษณ์ $P(X<9)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X<9) &= P(X=0)+\ldots+P(X=8) \\ &= 1-\{P(X=9)+P(X=10)\} \\ &= 1-P(X \geq 9) \\ &= 1-0.0017 \\ &= 0.9983 \end{aligned}$$

ความน่าจะเป็นที่มีผู้หายจากโรคนี้ไม่ถึง 9 คน มีค่าเท่ากับ 0.9983

ให้ Y เป็นตัวแปรสุ่มแทนจำนวนคนไข้ที่ไม่หายจากโรคนี้หลังได้รับการรักษา ทำให้สิ่งที่สนใจเปลี่ยนเป็นคนไข้ที่ไม่หายจากโรคนี้หลังได้รับการรักษา จากโจทย์จะได้ว่า $n=10$ และ p เท่ากับความน่าจะเป็นที่คนไข้จะไม่หายจากการเป็นโรคนี้หลังได้รับการรักษา นั่นคือ $p=0.6$ ดังนั้น

$$Y \sim B(10, 0.6)$$

การหาค่าความน่าจะเป็นของเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม Y สามารถหาได้โดยใช้ตารางที่ 1.10 ในภาคผนวก แต่เนื่องจากตารางไม่มีกรณีที่ $p > 0.5$ เราจึงจำเป็นต้องใช้ตัวแปรสุ่ม X เข้ามาช่วยในการหาค่าความน่าจะเป็นของตัวแปรสุ่ม Y

เนื่องจากเราสุ่มคนไข้มาทั้งหมด 10 คน ดังนั้น เหตุการณ์ที่สุ่มได้คนไข้ที่หายจากโรค k คน จะเป็นเหตุการณ์เดียวกับเหตุการณ์ที่สุ่มได้คนไข้ที่ไม่หายจากโรค $10-k$ คน ดังนั้น $P(Y=k)=P(X=10-k)$

(4) ความน่าจะเป็นที่มีผู้ไม่หายจากโรคนี้น้อยกว่า 1 คน แทนได้ด้วยสัญลักษณ์ $P(Y<1)$ ซึ่งหาค่าได้ดังนี้

$$P(Y<1)=P(Y=0)=P(X=10)=0.0001$$

ดังนั้นความน่าจะเป็นที่มีผู้ไม่หายจากโรคนี้น้อยกว่า 1 คน มีค่าเท่ากับ 0.0001

(5) ความน่าจะเป็นที่มีผู้ไม่หายจากโรคนี้ตั้งแต่ 8 คนขึ้นไป แทนได้ด้วยสัญลักษณ์ $P(Y \geq 8)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(Y \geq 8) &= P(Y=8)+P(Y=9)+P(Y=10) \\ &= P(X=2)+P(X=1)+P(X=0) \\ &= 0.1209+0.0403+0.0060 \\ &= 0.1672 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่มีผู้ไม่หายจากโรคนี้ตั้งแต่ 8 คนขึ้นไป มีค่าเท่ากับ 0.1672 ■

<u>ข้อสังเกต</u>

1. จากตัวอย่างที่ 3.3.4 ข้อ (2) จะสังเกตว่าเราจะเปิดตารางเพื่อหาค่า $P(X=0),\ \ldots,\ P(X=9)$ โดยตรงเลยก็ได้ หรือจะใช้สมบัติของความน่าจะเป็นแล้วเปิดตารางเพียง $P(X=10)$ ก็ได้ ซึ่งจะเห็นว่าวิธีที่สองจะทำให้เราหาค่าความน่าจะเป็นได้เร็วกว่าวิธีแรก

2. จากตัวอย่างที่ 3.3.4 ข้อ (4) และ (5) จะสังเกตได้ว่าเราไม่สามารถเปิดตารางการแจกแจงความน่าจะเป็นแบบทวินามได้ เนื่องจากตารางไม่มีกรณีที่ $p > 0.5$ ดังนั้นเราจึงจำเป็นต้องใช้ตัวแปรสุ่ม X เข้ามาช่วยในการหาค่าความน่าจะเป็นของตัวแปรสุ่ม Y (หรืออาจใช้ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทวินามในการคำนวณหาค่าความน่าจะเป็นแทนการเปิดตารางก็ได้)

3.4 การแจกแจงความน่าจะเป็นแบบปัวซง

<u>นิยามที่ 3.4.1</u> ตัวแปรสุ่มที่แทนจำนวนครั้งของเหตุการณ์หรือจำนวนผลลัพธ์ที่สนใจที่เกิดขึ้นในช่วงเวลาหรือขอบเขตที่กำหนด จะเรียกว่าเป็นตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบปัวซง (Poisson probability distribution)*** และเขียนแทนตัวแปรสุ่ม X ที่มีการแจกแจงความน่าจะเป็นแบบปัวซง ได้ด้วยสัญลักษณ์

$$X \sim Poi(\lambda)$$

และ λ แทนจำนวนผลลัพธ์ที่เกิดขึ้นเฉลี่ยในช่วงเวลาหรือขอบเขตที่กำหนด

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปัวซง

พิจารณา $X \sim Poi(\lambda)$ ดังนั้น ฟังก์ชันความน่าจะเป็น $f(x,\lambda)$ หาได้จากสมการ

$$f(x,\lambda) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

โดยที่ $x = 0,1,\ldots$ และ λ แทนจำนวนผลลัพธ์ที่เกิดขึ้นเฉลี่ยในช่วงเวลาหรือขอบเขตที่กำหนด

<u>หมายเหตุ</u>

ในหัวข้อนี้ใช้สัญลักษณ์ $f(x,\lambda)$ แทน $f(x)$ เนื่องจากจะแสดงให้เห็นว่าฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปัวซงมีความสัมพันธ์กับจำนวนผลลัพธ์ที่สนใจเฉลี่ยในขอบเขตที่กำหนด

ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปัวซง

พิจารณา $X \sim Poi(\lambda)$ จากนิยามค่าคาดหมายหรือค่าเฉลี่ยของตัวแปรสุ่มแบบไม่ต่อเนื่องจะได้ว่า

$$\begin{aligned}
E(X) &= \sum_x x f(x,\lambda) \\
&= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!} \\
&= e^{-\lambda}\sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} \\
&= e^{-\lambda}\sum_{x=1}^{\infty} x \frac{\lambda^x}{x!} \\
&= \lambda e^{-\lambda}\sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\ldots\right) \\
&= \lambda e^{-\lambda} e^{\lambda} = \lambda
\end{aligned}$$

(1) ความน่าจะเป็นที่จะมีลูกค้าเข้าร้านตั้งแต่ 3 คนขึ้นไปในแต่ละวัน แทนได้ด้วยสัญลักษณ์ $P(X \geq 3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X \geq 3) &= P(X=3)+P(X=4)+\ldots \\ &= 1-\{P(X=0)+P(X=1)+P(X=2)\} \\ &= 1-\left\{\frac{e^{-8}8^0}{0!}+\frac{e^{-8}8^1}{1!}+\frac{e^{-8}8^2}{2!}\right\} \\ &= 1-\{0.0003+0.0027+0.0107\} \\ &= 0.9863 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่จะมีลูกค้าเข้าร้านตั้งแต่ 3 คนขึ้นไปในแต่ละวันมีค่าเท่ากับ 0.9863

(2) ความน่าจะเป็นที่จะมีลูกค้าเข้าร้านน้อยกว่า 3 คนในแต่ละวัน แทนได้ด้วยสัญลักษณ์ $P(X<3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X<3) &= P(X=0)+P(X=1)+P(X=2) \\ &= \frac{e^{-8}8^0}{0!}+\frac{e^{-8}8^1}{1!}+\frac{e^{-8}8^2}{2!} \\ &= 0.0003+0.0027+0.0107 \\ &= 0.0137 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่จะมีลูกค้าเข้าร้านน้อยกว่า 3 คนในแต่ละวันมีค่าเท่ากับ 0.0137

(3) กำหนดให้ Y เป็นตัวแปรสุ่มแทนจำนวนลูกค้าที่เข้าร้านในระยะเวลาสองวัน จากโจทย์ทราบว่าโดยเฉลี่ยมีลูกค้าเข้าร้านวันละ 8 คน ดังนั้น λ ซึ่งมีค่าเท่ากับค่าเฉลี่ยของจำนวนลูกค้าที่เข้าร้านในระยะเวลาสองวันจึงมีค่าเท่ากับ 16 ดังนั้น

$$Y \sim Poi(16)$$

ความน่าจะเป็นที่จะมีลูกค้าเข้าร้านน้อยกว่า 3 คนในระยะเวลาสองวัน แทนได้ด้วยสัญลักษณ์ $P(Y<3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(Y<3) &= P(Y=0)+P(Y=1)+P(Y=2) \\ &= \frac{e^{-16}(16)^0}{0!}+\frac{e^{-16}(16)^1}{1!}+\frac{e^{-16}(16)^2}{2!} \\ &= 0+0+0=0 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่จะมีลูกค้าเข้าร้านน้อยกว่า 3 คนในระยะเวลาสองวันมีค่าเท่ากับ 0

(4) กำหนดให้ Z เป็นตัวแปรสุ่มแทนจำนวนลูกค้าที่เข้าร้านในระยะเวลาหนึ่งวันครึ่ง จากโจทย์ทราบว่าโดยเฉลี่ยมีลูกค้าเข้าร้านวันละ 8 คน ดังนั้น λ ซึ่งมีค่าเท่ากับค่าเฉลี่ยของจำนวนลูกค้าที่เข้าร้านในระยะเวลาหนึ่งวันครึ่ง จึงมีค่าเท่ากับ 12 ดังนั้น

$$Z \sim Poi(12)$$

ความน่าจะเป็นที่จะมีลูกค้าเข้าร้านอย่างน้อย 3 คนในระยะเวลาหนึ่งวันครึ่ง แทนได้ด้วยสัญลักษณ์ $P(Z \geq 3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(Z \geq 3) &= P(Z=3)+P(X=4)+\ldots \\ &= 1-\{P(Z=0)+P(X=1)+P(Z=2)\} \\ &= 1-\left\{\frac{e^{-12}(12)^0}{0!}+\frac{e^{-12}(12)^1}{1!}+\frac{e^{-12}(12)^2}{2!}\right\} \\ &= 1-\{0+0.0001+0.0004\} \\ &= 1-0.0005=0.9995 \end{aligned}$$

ดังนั้น ความน่าจะเป็นที่จะมีลูกค้าเข้าร้านอย่างน้อย 3 คนในระยะเวลาหนึ่งวันครึ่งมีค่าเท่ากับ 0.9995

(5) กำหนดให้ A เป็นตัวแปรสุ่มแทนจำนวนลูกค้าที่เข้าร้านในระยะเวลา 4 ชั่วโมง จากโจทย์ทราบว่าร้านขายรถยนต์แห่งนี้เปิดทำการวันละ 8 ชั่วโมงต่อวัน และโดยเฉลี่ยมีลูกค้าเข้าร้านวันละ 8 คน (นั่นคือ 8 ชั่วโมงลูกค้าเข้าร้านโดยเฉลี่ย 8 คน) ดังนั้น λ ซึ่งมีค่าเท่ากับค่าเฉลี่ยของจำนวนลูกค้าที่เข้าร้านในระยะเวลา 4 ชั่วโมงจึงมีค่าเท่ากับ 4 ดังนั้น

$$A \sim Poi(4)$$

ความน่าจะเป็นที่จะมีลูกค้าเข้าร้านไม่เกิน 3 คน ในระยะเวลา 4 ชั่วโมง แทนได้ด้วยสัญลักษณ์ $P(A \leq 3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(A \leq 3) &= P(A=0)+P(A=1)+P(A=2)+P(A=3) \\ &= \frac{e^{-4}(4)^0}{0!}+\frac{e^{-4}(4)^1}{1!}+\frac{e^{-4}(4)^2}{2!}+\frac{e^{-4}(4)^3}{3!} \\ &= 0.0183+0.0733+0.1465+0.1954 \\ &= 0.4335 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่จะมีลูกค้าเข้าร้านไม่เกิน 3 คน ในระยะเวลา 4 ชั่วโมง มีค่าเท่ากับ 0.4335 ■

<u>ข้อสังเกต</u>

1. จากตัวอย่างที่ 3.4.2 ข้อ (3) จะสังเกตได้ว่าคำถามนั้นถามเกี่ยวกับจำนวนลูกค้าที่เข้าร้านในระยะเวลาสองวัน ซึ่งแตกต่างกับข้อ (1) และ (2) ที่ถามเกี่ยวกับจำนวนลูกค้าที่เข้าร้านในระยะเวลาหนึ่งวัน ดังนั้นตัวแปรสุ่มในข้อ (3) จึงแตกต่างจากตัวแปรสุ่มในข้อ (1) และ (2) ในทำนองเดียวกัน ตัวอย่างที่ 3.4.2 ข้อ (4) และ (5) ก็เช่นกัน

2. การคำนวณหาค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปัวซง สามารถหาค่าได้โดยตรงจากฟังก์ชันความน่าจะเป็น หรือใช้ตารางที่ 2 ในภาคผนวก

<u>ตัวอย่างที่ 3.4.3</u> พนักงานคนหนึ่ง โดยเฉลี่ยจะรับโทรศัพท์วันละ 14 ครั้ง จงหา

(1) ความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์ไม่น้อยกว่า 10 ครั้ง ในแต่ละวัน

(2) ถ้าใน 1 วัน พนักงานคนนี้ทำงานทั้งหมด 10 ชั่วโมง จงหาความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์มากกว่า 2 ครั้ง ในเวลา 1 ชั่วโมง

(3) ถ้าใน 1 เดือน พนักงานคนนี้มาทำงาน 20 วัน จงหาความน่าจะเป็นที่จะรับโทรศัพท์ไม่น้อยกว่า 10 ครั้งต่อวัน เป็นจำนวนวันทั้งสิ้น 15 วัน

วิธีทำ (1) กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนครั้งของการรับโทรศัพท์ในแต่ละวัน เนื่องจากตัวแปรสุ่ม X ที่กำหนดเป็นการนับจำนวนของผลลัพธ์ที่เกิดขึ้นในช่วงเวลา 1 วัน และจากโจทย์ λ มีค่าเท่ากับค่าเฉลี่ยของจำนวนครั้งที่พนักงานคนนี้รับโทรศัพท์ในแต่ละวัน นั่นคือ $\lambda = 14$ ดังนั้น

$$X \sim Poi(14)$$

ดังนั้น การหาค่าความน่าจะเป็นของเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม X สามารถหาได้โดยใช้ตารางที่ 2.11 ในภาคผนวก

ความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์ไม่น้อยกว่า 10 ครั้ง ในแต่ละวัน แทนได้ด้วยสัญลักษณ์ $P(X \geq 10)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned}
P(X \geq 10) &= P(X = 10) + P(X = 11) + \ldots \\
&= 1 - \{P(X = 0) + \ldots + P(X = 9)\} \\
&= 1 - \{0 + 0 + 0.0001 + 0.0004 + 0.0013 + 0.0037 + 0.0087 + 0.0174 + 0.0304 + 0.0473\} \\
&= 1 - 0.1093 \\
&= 0.8907
\end{aligned}$$

ดังนั้น ความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์ไม่น้อยกว่า 10 ครั้ง ในแต่ละวันมีค่าเท่ากับ 0.8907

(2) กำหนดให้ Y เป็นตัวแปรสุ่มแทนจำนวนครั้งของการรับโทรศัพท์ในเวลา 1 ชั่วโมง จากโจทย์ทราบว่าพนักงานคนนี้ทำงานทั้งหมด 10 ชั่วโมงต่อวัน และโดยเฉลี่ยพนักงานคนนี้จะรับโทรศัพท์วันละ 14 ครั้ง (นั่นคือ 10 ชั่วโมง รับโทรศัพท์ 14 ครั้ง) ดังนั้น λ ซึ่งมีค่าเท่ากับค่าเฉลี่ยของจำนวนครั้งที่พนักงานคนนี้รับโทรศัพท์ใน 1 ชั่วโมง จึงมีค่าเท่ากับ 1.4 ดังนั้น

$$Y \sim Poi(1.4)$$

ดังนั้น การหาค่าความน่าจะเป็นของเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม X สามารถหาได้โดยใช้ตารางที่ 2.2 ในภาคผนวก

ความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์มากกว่า 2 ครั้ง ในเวลา 1 ชั่วโมง แทนได้ด้วยสัญลักษณ์ $P(Y > 2)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned}
P(Y > 2) &= P(Y = 3) + \ldots \\
&= 1 - \{P(Y = 0) + P(Y = 1) + P(Y = 2)\} \\
&= 1 - \{0.2466 + 0.3452 + 0.2417\} \\
&= 1 - 0.8335 \\
&= 0.1665
\end{aligned}$$

ดังนั้น ความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์มากกว่า 2 ครั้ง ในเวลา 1 ชั่วโมง มีค่าเท่ากับ 0.1665

(3) พิจารณาการทดลองสุ่มเลือกวันมา 1 วัน ผลลัพธ์ที่ได้มีเพียง 2 แบบ เท่านั้น คือ สุ่มได้วันที่พนักงานคนนี้รับโทรศัพท์ไม่น้อยกว่า 10 ครั้งต่อวัน หรือสุ่มได้วันที่รับโทรศัพท์ตั้งแต่ 10 ครั้งต่อวัน สิ่งที่สนใจคือการสุ่มได้วันที่รับโทรศัพท์ไม่น้อยกว่า 10 ครั้งต่อวัน ซึ่งมีความน่าจะเป็นคงที่เสมอ นั่นคือการสุ่มเลือกวันมา 1 วัน เป็นการทดลองแบบเบอร์นูลี ดังนั้นการสุ่มเลือกวันมา 20 วัน ถือว่าเป็นการทดลองแบบเบอร์นูลีซ้ำกัน 20 ครั้ง และการสุ่มเลือกวันแต่ละวันเป็นอิสระต่อกัน ดังนั้นจึงถือว่าเป็นการทดลองแบบทวินาม

กำหนดให้ Z เป็นตัวแปรสุ่มแทนจำนวนวันที่รับโทรศัพท์ไม่น้อยกว่า 10 ครั้งต่อวัน จากโจทย์จะได้ว่า $n=20$ และ p เท่ากับความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์ไม่น้อยกว่า 10 ครั้งต่อวัน ซึ่งหาค่ามาแล้วจากข้อ (1) นั่นคือ $p=0.8907$ ดังนั้น

$$Z \sim B(20, 0.8907)$$

ความน่าจะเป็นที่พนักงานคนนี้จะรับโทรศัพท์ไม่น้อยกว่า 10 ครั้งต่อวัน เป็นจำนวนวันทั้งสิ้น 15 วัน แทนได้ด้วยสัญลักษณ์ $P(Z=15)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(Z=15) &= \binom{20}{15}(0.8907)^{15}(1-0.8907)^{20-15} \\ &= 0.0426 \end{aligned}$$

ดังนั้น ถ้าใน 1 เดือน พนักงานคนนี้มาทำงาน 20 วัน ความน่าจะเป็นที่จะรับโทรศัพท์ไม่น้อยกว่า 10 ครั้งต่อวัน เป็นจำนวนวันทั้งสิ้น 15 วัน มีค่าเท่ากับ 0.0426 ■

<u>ข้อสังเกต</u>

ตัวอย่างที่ 3.4.3 ข้อ (3) เป็นตัวอย่างที่สัมพันธ์กับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทวินามและแบบปัวซง

3.4.1 การประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปัวซง

กำหนดตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบทวินาม โดยที่ $n \to \infty,\ p \to 0$ และ np มีค่าคงที่ กำหนดให้ $\lambda = np$ หรือ $p = \dfrac{\lambda}{n}$ พิจารณา

$$\begin{aligned} \lim_{n\to\infty}\binom{n}{x}p^x(1-p)^{n-x} &= \lim_{n\to\infty}\frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} \\ &= \lim_{n\to\infty}\frac{n(n-1)\ldots(n-x+1)}{x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x} \\ &= \frac{\lambda^x}{x!}\lim_{n\to\infty}\frac{n(n-1)\ldots(n-x+1)}{n^x}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-x} \end{aligned}$$

วิธีทำ พิจารณาการทดลองสุ่มเด็กมา 1 คน ผลลัพธ์ที่ได้มีเพียง 2 แบบ เท่านั้น คือ เด็กแพ้วัคซีน หรือเด็กไม่แพ้วัคซีน สิ่งที่สนใจคือการสุ่มเจอเด็กที่แพ้วัคซีน ซึ่งมีความน่าจะเป็นคงที่เสมอ นั่นคือการสุ่มเด็กมา 1 คน เป็นการทดลองแบบเบอร์นูลี ดังนั้นการสุ่มเด็กมา 2,000 คน ถือว่าเป็นการทดลองแบบเบอร์นูลีซ้ำกัน 2,000 ครั้ง ถ้าเป็นการสุ่มแบบใส่คืน การทดลองแต่ละครั้งเป็นอิสระต่อกัน จะถือว่าเป็นการทดลองแบบทวินาม สำหรับโจทย์ข้อนี้เป็นการสุ่มแบบไม่ใส่คืน แต่เนื่องจากสุ่มจากเด็กทั่วไปซึ่งถือว่ามีจำนวนมาก ดังนั้นจึงถือว่าเป็นการทดลองแบบทวินาม

ในตัวอย่างนี้จะแสดงวิธีหาค่าความน่าจะเป็น 2 วิธี ดังนี้

วิธีที่ 1 คำนวณโดยตรงจากฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม

กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนเด็กที่แพ้วัคซีน ดังนั้น

$$X \sim B(2000, 0.0001)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการ

$$f(x, 2000, 0.0001) = \binom{2000}{x}(0.0001)^x(1-0.0001)^{2000-x}$$

(1) ความน่าจะเป็นที่เด็กจะแพ้วัคซีนจำนวน 3 คน แทนได้ด้วยสัญลักษณ์ $P(X=3)$ ซึ่งหาค่าได้ดังนี้

$$P(X=3) = \binom{2000}{3}(0.0001)^3(1-0.0001)^{2000-3} = 0.0011$$

ดังนั้นความน่าจะเป็นที่เด็กจะแพ้วัคซีนจำนวน 3 คน มีค่าเท่ากับ 0.0011

(2) ความน่าจะเป็นที่เด็กจะแพ้วัคซีนมากกว่า 3 คน แทนได้ด้วยสัญลักษณ์ $P(X>3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned}
P(X>3) &= P(X=4)+\ldots+P(X=2000) \\
&= 1-\{P(X=0)+\ldots+P(X=3)\} \\
&= 1-\left\{\binom{2000}{0}(0.0001)^0(1-0.0001)^{2000-0}+\ldots+\binom{2000}{3}(0.0001)^3(1-0.0001)^{2000-3}\right\} \\
&= 1-\{0.8187+0.1638+0.0164+0.0011\} \\
&= 1-1=0
\end{aligned}$$

ดังนั้นความน่าจะเป็นที่เด็กจะแพ้วัคซีนมากกว่า 3 คน มีค่าเท่ากับ 0

กำหนดให้ Y เป็นตัวแปรสุ่มแทนจำนวนเด็กที่ไม่แพ้วัคซีน ดังนั้น

$$Y \sim B(2000, 0.9999)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการ

$$f(y, 2000, 0.9999) = \binom{2000}{y}(0.9999)^y(1-0.9999)^{2000-y}$$

(3) ความน่าจะเป็นที่เด็กไม่แพ้วัคซีนน้อยกว่า 1998 คน แทนได้ด้วยสัญลักษณ์ $P(Y<1998)$ ซึ่งหาค่าได้ดังนี้

$$P(Y<1998) = P(Y=0)+\ldots+P(Y=1997)$$

$$= 1-\{P(Y=1998)+P(Y=1999)+P(Y=2000)\}$$
$$= 1-\left\{\binom{2000}{1998}(0.9999)^{1998}(1-0.9999)^{2000-1998}+\ldots+\binom{2000}{2000}(0.9999)^{2000}(1-0.9999)^{2000-2000}\right\}$$
$$= 1-\{0.0164+0.1638+0.8187\}$$
$$= 1-0.9989 = 0.0011$$

ดังนั้นความน่าจะเป็นที่เด็กไม่แพ้วัคซีนน้อยกว่า 1998 คน มีค่าเท่ากับ 0.0011

(4) ความน่าจะเป็นที่เด็กไม่แพ้วัคซีนตั้งแต่ 1995 คน แทนได้ด้วยสัญลักษณ์ $P(Y \geq 1995)$ ซึ่งหาค่าได้ดังนี้

$$P(Y \geq 1995) = P(Y=1995)+\ldots+P(Y=2000)$$
$$= \binom{2000}{1995}(0.9999)^{1995}(1-0.9999)^{2000-1995}+\ldots+\binom{2000}{2000}(0.9999)^{2000}(1-0.9999)^{2000-2000}$$
$$= 0+0.0001+0.0011+0.0164+0.1638+0.8187 = 1$$

ดังนั้นความน่าจะเป็นที่เด็กไม่แพ้วัคซีนตั้งแต่ 1995 คน มีค่าเท่ากับ 1

<u>วิธีที่ 2</u> คำนวณโดยใช้การประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปัวซง

กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนเด็กที่แพ้วัคซีน ดังนั้น

$$X \sim B(2000, 0.0001)$$

เนื่องจาก $n = 2000$ มีค่ามาก และ $p = 0.0001$ มีค่าน้อยเข้าใกล้ศูนย์ ดังนั้นเราสามารถประมาณค่าความน่าจะเป็นแบบทวินามได้ด้วยความน่าจะเป็นแบบปัวซง โดยที่ $\lambda = np = (2000)(0.0001) = 0.2$ ซึ่งการหาค่าความน่าจะเป็นของเหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่ม X สามารถหาได้โดยใช้ตารางที่ 2.1 ในภาคผนวก

(1) ความน่าจะเป็นที่เด็กจะแพ้วัคซีนจำนวน 3 คน แทนได้ด้วยสัญลักษณ์ $P(X=3)$ ซึ่งหาค่าได้ดังนี้

$$P(X=3) = 0.0011$$

ดังนั้นความน่าจะเป็นที่เด็กจะแพ้วัคซีนจำนวน 3 คน มีค่าเท่ากับ 0.0011

(2) ความน่าจะเป็นที่เด็กจะแพ้วัคซีนมากกว่า 3 คน แทนได้ด้วยสัญลักษณ์ $P(X>3)$ ซึ่งหาค่าได้ดังนี้

$$P(X>3) = P(X=4)+\ldots+P(X=2000)$$
$$= 1-\{P(X=0)+\ldots+P(X=3)\}$$
$$= 1-\{0.8187+0.1637+0.0164+0.0011\}$$
$$= 1-0.9999 = 0.0001$$

ดังนั้นความน่าจะเป็นที่เด็กจะแพ้วัคซีนมากกว่า 3 คน มีค่าเท่ากับ 0.0001

ให้ Y เป็นตัวแปรสุ่มแทนจำนวนเด็กที่ไม่แพ้วัคซีน ดังนั้น

$$Y \sim B(2000, 0.9999)$$

แต่เนื่องจาก $p = 0.9999$ มีค่ามากไม่เข้าใกล้ศูนย์ จึงไม่สามารถใช้การประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปัวซงได้ ดังนั้น จึงใช้ตัวแปรสุ่ม X เข้ามาช่วยในการหาค่าความน่าจะเป็นแทน โดยที่ $P(Y = k) = P(X = 2000 - k)$

(3) ความน่าจะเป็นที่เด็กไม่แพ้วัคซีนน้อยกว่า 1998 คน แทนได้ด้วยสัญลักษณ์ $P(Y < 1998)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(Y < 1998) &= P(Y = 0) + \ldots + P(Y = 1997) \\ &= P(X = 2000) + \ldots + P(X = 3) \\ &= 1 - \{P(X = 0) + P(X = 1) + P(X = 2)\} \\ &= 1 - \{0.8187 + 0.1637 + 0.0164\} \\ &= 1 - 0.9988 = 0.0012 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่เด็กไม่แพ้วัคซีนน้อยกว่า 1998 คน มีค่าเท่ากับ 0.0012

(4) ความน่าจะเป็นที่เด็กไม่แพ้วัคซีนตั้งแต่ 1995 คน แทนได้ด้วยสัญลักษณ์ $P(Y \geq 1995)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(Y \geq 1995) &= P(Y = 1995) + \ldots + P(Y = 2000) \\ &= P(X = 5) + \ldots + P(X = 0) \\ &= 0 + 0.0001 + 0.0011 + 0.0164 + 0.1637 + 0.8187 = 1 \end{aligned}$$

ดังนั้นความน่าจะเป็นที่เด็กไม่แพ้วัคซีนตั้งแต่ 1995 คน มีค่าเท่ากับ 1 ■

<u>หมายเหตุ</u>

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบไม่ต่อเนื่องนอกเหนือจากที่ได้กล่าวมาแล้วในบทนี้ยังมีอีกหลายชนิด เช่น การแจกแจงความน่าจะเป็นแบบพหุนาม (multinomial probability distribution) การแจกแจงความน่าจะเป็นแบบเรขาคณิต (geometric probability distribution) การแจกแจงความน่าจะเป็นแบบทวินามลบ (negative binomial) การแจกแจงความน่าจะเป็นแบบไฮเพอร์จีออเมตริก (hypergeometric probability distribution) เป็นต้น ซึ่งสามารถหาอ่านเพิ่มเติมได้ในหนังสือเกี่ยวกับความน่าจะเป็นและสถิติ

บทที่ 4
การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่อง
(Continuous Probability Distributions)

ในบทที่ 4 นี้จะกล่าวถึงการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องที่สำคัญ ได้แก่ การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง การแจกแจงความน่าจะเป็นแบบปกติ

4.1 การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องที่ง่ายที่สุด คือ การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง ซึ่งจะได้กล่าวถึงในหัวข้อนี้

นิยามที่ 4.1.1 ถ้า X เป็นตัวแปรสุ่มแบบต่อเนื่องมีค่าที่เป็นไปได้ทั้งหมดอยู่ในช่วง $[a,b]$ แล้ว จะเรียก X ว่าเป็นตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง (continuous uniform probability distribution)*** เมื่อตัวแปรสุ่ม X มีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x,a,b)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b]\\ 0, & x\notin[a,b]\end{cases}$$

ถ้าตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่องและมีค่าที่เป็นไปได้อยู่ในช่วง $[a,b]$ จะเขียนแทนได้ด้วยสัญลักษณ์

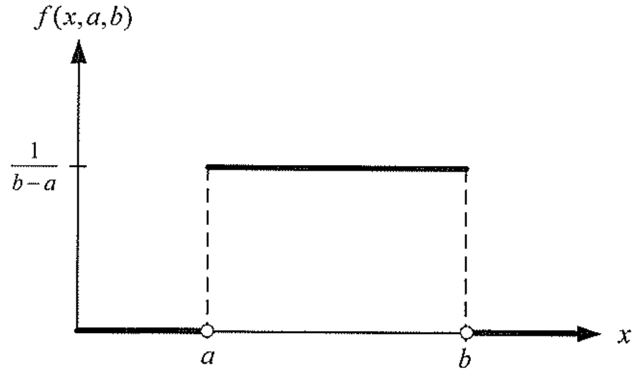
$$X\sim U(a,b)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง

พิจารณา $X\sim U(a,b)$ จากนิยามที่ 4.1.1 ดังนั้น ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบยูนิฟอร์มต่อเนื่อง เป็นไปตามสมการ

$$f(x,a,b)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b]\\ 0, & x\notin[a,b]\end{cases}$$

สำหรับฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่องแสดงได้ในรูปที่ 4.1.1



รูปที่ 4.1.1 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง

<u>หมายเหตุ</u>

ในหัวข้อนี้ใช้สัญลักษณ์ $f(x,a,b)$ แทน $f(x)$ เนื่องจากจะแสดงให้เห็นว่าฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่องมีความสัมพันธ์กับค่า a และ b

ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง

พิจารณา $X\sim U(a,b)$ จากนิยามค่าคาดหมายหรือค่าเฉลี่ยของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$\begin{aligned}E(X) &= \int_{-\infty}^{\infty} xf(x,a,b)dx\\ &= \int_a^b x\cdot\frac{1}{b-a}dx\\ &= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_{x=a}^{x=b}\\ &= \frac{b^2-a^2}{2(b-a)}\\ &= \frac{a+b}{2}\end{aligned}$$

ดังนั้น ค่าเฉลี่ยหรือค่าคาดหมายของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่องมีค่าเท่ากับ

$$E(X)=\frac{a+b}{2}$$

ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง

พิจารณา $X \sim U(a,b)$ จากนิยามความแปรปรวนของตัวแปรสุ่ม $V(X)=E(X^2)-(E(X))^2$ ดังนั้น จึงหาค่า $E(X^2)$ ได้ดังนี้

$$\begin{aligned}E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx \\ &= \int_a^b x^2 \cdot \frac{1}{b-a}dx \\ &= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_{x=a}^{x=b} \\ &= \frac{b^3-a^3}{3(b-a)} \\ &= \frac{b^2+ab+a^2}{3}\end{aligned}$$

แทนค่า $E(X^2)$ และ $E(X)$ เพื่อหาความแปรปรวนของตัวแปรสุ่ม X ได้ดังนี้

$$\begin{aligned}V(X) &= E(X^2)-(E(X))^2 \\ &= \frac{b^2+ab+a^2}{3}-\left(\frac{a+b}{2}\right)^2 \\ &= \frac{b^2+ab+a^2}{3}-\frac{a^2+2ab+b^2}{4} \\ &= \frac{b^2-2ab+a^2}{12} \\ &= \frac{(b-a)^2}{12}\end{aligned}$$

ดังนั้น ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่องมีค่าเท่ากับ

$$V(X)=\frac{(b-a)^2}{12}$$

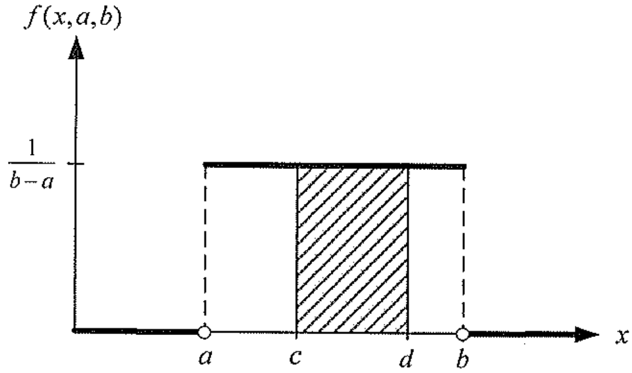
<u>ข้อสังเกต</u>

1. พิจารณา $X \sim U(a,b)$ ดังนั้น $P(c \le X \le d)$ โดยที่ $a \le c \le d \le b$ สามารถหาค่าได้ดังนี้

$$P(c \le X \le d)=\int_c^d f(x,a,b)dx=\int_c^d \frac{1}{b-a}dx=\frac{1}{b-a}[x]_{x=c}^{x=d}=\frac{d-c}{b-a}$$

เมื่อพิจารณาพื้นที่ใต้กราฟของ $f(x)$ เมื่อ $a \le c \le x \le d \le b$ ซึ่งเป็นพื้นที่บริเวณที่แรเงา ดังรูปที่ 4.1.2 พบว่ามีพื้นที่เท่ากับ $\dfrac{d-c}{b-a}$ ซึ่งมีค่าเท่ากับ $P(c \le X \le d)$

ดังนั้น การหาค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง สามารถหาค่าได้จากพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นในช่วงที่ต้องการนั่นเอง



รูปที่ 4.1.2 พื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง

2. จากข้อสังเกตข้อที่ 1 ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง จะมีความน่าจะเป็นเท่ากันในแต่ละช่วงที่มีความยาวเท่ากัน (กำหนดให้ความยาวของช่วงมีค่าเท่ากับ δ) ซึ่งอยู่ระหว่าง a และ b และมีความน่าจะเป็นเท่ากับ $\dfrac{\delta}{b-a}$

3. สำหรับ $X \sim (0,1)$ จะเรียกตัวแปรสุ่ม X ว่าเป็นตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบยูนิฟอร์มมาตรฐานต่อเนื่อง (standard uniform probability distribution)***

สำหรับ $X \sim U(a,b)$ สามารถแปลงเป็นตัวแปรสุ่ม $Y \sim U(0,1)$ ได้ตามความสัมพันธ์ต่อไปนี้

$$Y = \frac{X-a}{b-a} \sim U(0,1)$$

<u>**ตัวอย่างที่ 4.1.1**</u> กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนจริงที่มีค่าที่เป็นไปได้อยู่ในช่วง $[10,30]$ ซึ่งมีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง จงหา

(1) ความน่าจะเป็นที่ตัวแปรสุ่ม X จะมีค่าตั้งแต่ 10 ถึง 15
(2) ความน่าจะเป็นที่ตัวแปรสุ่ม X จะมีค่ามากกว่า 35
(3) ความน่าจะเป็นที่ตัวแปรสุ่ม X จะมีค่าน้อยกว่าหรือเท่ากับ 14
(4) ความน่าจะเป็นที่ตัวแปรสุ่ม X จะมีค่ามากกว่าหรือเท่ากับ 20
(5) ค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่ม X

วิธีทำ เนื่องจากโจทย์กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนจริงที่มีค่าที่เป็นไปได้อยู่ในช่วง $[10,30]$ ซึ่งมีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์มต่อเนื่อง ดังนั้น

$$X \sim U(10,30)$$

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X คือ

$$f(x,10,30)=\begin{cases}\frac{1}{20}, & x\in[10,30]\\ 0, & x\notin[10,30]\end{cases}$$

(1) ความน่าจะเป็นที่ X จะมีค่าตั้งแต่ 10 ถึง 15 แทนได้ด้วยสัญลักษณ์ $P(10\le X\le 15)$ ซึ่งหาค่าได้ดังนี้

$$P(10\le X\le 15)=\int_{10}^{15} f(x,10,30)dx=\int_{10}^{15}\frac{1}{20}dx=\frac{1}{20}\left[x\right]_{x=10}^{x=15}=0.25$$

ดังนั้น ตัวแปรสุ่ม X มีค่าตั้งแต่ 10 ถึง 15 ด้วยความน่าจะเป็นเท่ากับ 0.25

(2) ความน่าจะเป็นที่ X จะมีค่ามากกว่า 35 แทนได้ด้วยสัญลักษณ์ $P(X>35)$ ซึ่งหาค่าได้ดังนี้

$$P(X>35)=\int_{35}^{\infty} f(x,10,30)dx=\int_{35}^{\infty}0dx=0$$

ดังนั้น ตัวแปรสุ่ม X มีค่ามากกว่า 35 ด้วยความน่าจะเป็นเท่ากับ 0

(3) ความน่าจะเป็นที่ X จะมีค่าน้อยกว่าหรือเท่ากับ 14 แทนได้ด้วยสัญลักษณ์ $P(X\le 14)$ ซึ่งหาค่าได้ดังนี้

$$P(X\le 14)=\int_{-\infty}^{14} f(x,10,30)dx=\int_{10}^{14}\frac{1}{20}dx=\frac{1}{20}\left[x\right]_{x=10}^{x=14}=0.2$$

ดังนั้น ตัวแปรสุ่ม X มีค่าน้อยกว่าหรือเท่ากับ 14 ด้วยความน่าจะเป็นเท่ากับ 0.2

(4) ความน่าจะเป็นที่ X จะมีค่ามากกว่าหรือเท่ากับ 20 แทนได้ด้วยสัญลักษณ์ $P(X\ge 20)$ ซึ่งหาค่าได้ดังนี้

$$P(X\ge 20)=\int_{20}^{\infty} f(x,10,30)dx=\int_{10}^{30}\frac{1}{20}dx=\frac{1}{20}\left[x\right]_{x=20}^{x=30}=0.5$$

ดังนั้น ตัวแปรสุ่ม X มีค่ามากกว่าหรือเท่ากับ 20 ด้วยความน่าจะเป็นเท่ากับ 0.5

(5) ค่าเฉลี่ยของตัวแปรสุ่ม X แทนได้ด้วยสัญลักษณ์ $E(X)$ ซึ่งหาค่าได้ดังนี้

$$E(X)=\frac{a+b}{2}=\frac{10+30}{2}=20$$

ความแปรปรวนของตัวแปรสุ่ม X แทนได้ด้วยสัญลักษณ์ $V(X)$ ซึ่งหาค่าได้ดังนี้

$$V(X)=\frac{(b-a)^2}{12}=\frac{(30-10)^2}{12}=\frac{100}{3}$$

ตามสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องที่เคยกล่าวไว้แล้วในบทที่ 2 นั่นคือ $\int_{-\infty}^{\infty} f(z,0,1)dz = 1$ ดังนั้น ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมีค่าเท่ากับ

$$E(X) = \mu$$

ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติ

พิจารณา $X \sim N(\mu, \sigma^2)$ จากนิยามความแปรปรวนของตัวแปรสุ่มจะได้ว่า

$$\begin{aligned} V(X) &= E\left((X - E(X))^2\right) \\ &= E\left((X-\mu)^2\right) \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x,\mu,\sigma)dx \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \end{aligned}$$

กำหนดให้ $z = \frac{x-\mu}{\sigma}$ จะได้ว่า $dz = \frac{1}{\sigma}dx$ ดังนั้น

$$\begin{aligned} V(X) &= \sigma \int_{-\infty}^{\infty} \frac{(\sigma z)^2}{\sigma\sqrt{2\pi}} e^{\frac{-z^2}{2}} dz \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{\frac{-z^2}{2}} dz \end{aligned}$$

พิจารณาการหาค่าด้วยการอินทิเกรตแบบแยกส่วน (by parts) โดยให้ $u = z$ และ $dv = ze^{\frac{-z^2}{2}}$ ดังนั้น

$$V(X) = \frac{\sigma^2}{\sqrt{2\pi}}\left(-ze^{\frac{-z^2}{2}}\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz \right)$$

เนื่องจาก $\lim_{z\to\infty} ze^{\frac{-z^2}{2}} = \lim_{z\to-\infty} ze^{\frac{-z^2}{2}} = 0$ ดังนั้น

$$\begin{aligned} V(X) &= \frac{\sigma^2}{\sqrt{2\pi}}\left(0 + \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz \right) \\ &= \sigma^2 \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz \end{aligned}$$

เนื่องจาก $f(z,0,1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-z^2}{2}} dz$ เป็นฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติ ดังนั้น

$$V(X) = \sigma^2 \int_{-\infty}^{\infty} f(z,0,1)dz$$

ตามสมบัติของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มแบบต่อเนื่องที่เคยกล่าวไว้แล้วในบทที่ 2 นั่นคือ $\int_{-\infty}^{\infty} f(z,0,1)dz = 1$ ดังนั้น ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมีค่าเท่ากับ

$$V(X) = \sigma^2$$

<u>หมายเหตุ</u>

ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ 0 ความแปรปรวนเท่ากับ 1 เรียกว่า ตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน (standard normal probability distribution)*** และกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน เรียกว่า ***โค้งปกติมาตรฐาน (standard normal curve)***

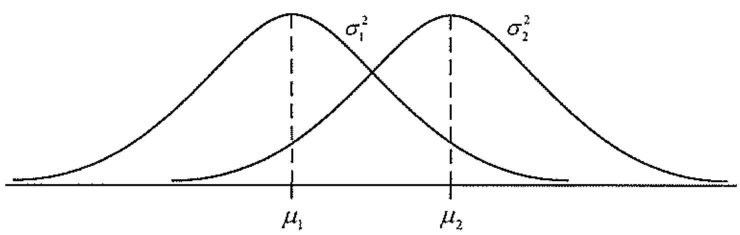
สมบัติที่สำคัญของโค้งปกติ

เมื่อกำหนดให้ $X \sim N(\mu,\sigma^2)$

1. โค้งปกติเป็นเส้นโค้งรูประฆังคว่ำ ตามรูปที่ 4.2.1
2. พื้นที่ใต้โค้งปกติแต่เหนือแกน X มีค่าเท่ากับ 1
3. โค้งปกติมีลักษณะสมมาตรตามแนวแกน Y ที่ $X = \mu$ โดยเส้นตรงนี้แบ่งพื้นที่ใต้โค้งปกติออกเป็น 2 ส่วน เท่า ๆ กัน ตามรูปที่ 4.2.1
4. โค้งปกติมีค่าสูงสุดเท่ากับ $\frac{1}{\sigma\sqrt{2\pi}}$ ซึ่งเกิดขึ้นเมื่อ $X = \mu$
5. โค้งปกติมีจุดเปลี่ยนเว้าที่จุด $X = \mu \pm \sigma$ โดยที่กราฟเว้าคว่ำเมื่อ $\mu - \sigma < X < \mu + \sigma$ และกราฟเว้าหงายเมื่อ $X < \mu - \sigma$ และ $X > \mu + \sigma$
6. โค้งปกติจะไม่ตัดแกน X แต่มีเส้นกำกับในแนวนอนคือแกน X

<u>หมายเหตุ</u>

1. เมื่อเปรียบเทียบโค้งปกติ 2 เส้น ซึ่งมีความแปรปรวนเท่ากัน แต่ค่าเฉลี่ยต่างกัน แสดงได้ในรูปที่ 4.2.2



รูปที่ 4.2.2 โค้งปกติ เมื่อ $\mu_1 < \mu_2$ และ $\sigma_1^2 = \sigma_2^2$

2. เมื่อเปรียบเทียบโค้งปกติ 2 เส้น ซึ่งมีค่าเฉลี่ยเท่ากัน แต่ความแปรปรวนต่างกัน แสดงได้ในรูปที่ 4.2.3

$$P(x_1 < X < x_2) = \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

ดังนั้น $P(x_1 < X < x_2)$ มีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $z_1 < Z < z_2$ นั่นเอง

<u>ตัวอย่างที่ 4.2.1</u> กำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน นั่นคือ $X \sim N(0,1)$ จงหา

(1) $P(X \le 2.02)$

(2) $P(X \ge 1.65)$

(3) $P(X < -1.05)$

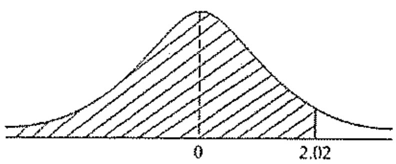
(4) $P(X > -1.89)$

(5) $P(1.11 \le X < 2.22)$

(6) $P(-1.06 \le X < -0.55)$

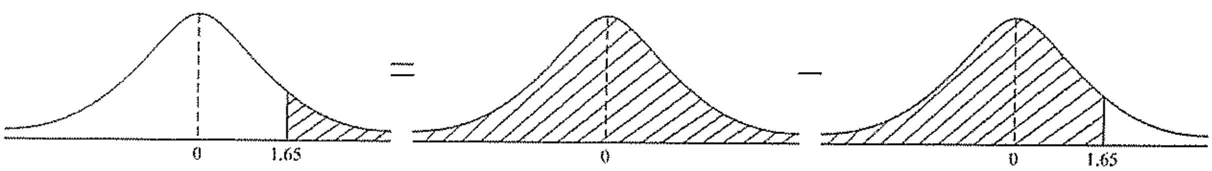
(7) $P(-2.01 \le X \le 1.34)$

วิธีทำ (1) พิจารณา $P(X \le 2.02)$ จะเห็นว่ามีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X \le 2.02$ ดังรูป



โดยใช้ตารางที่ 3.1 ได้ว่า $P(X \le 2.02) = 0.9783$ ดังนั้น $X \le 2.02$ ด้วยความน่าจะเป็นเท่ากับ 0.9783

(2) พิจารณา $P(X \ge 1.65)$ จะเห็นว่ามีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X \ge 1.65$ เนื่องจากพื้นที่ใต้โค้งปกติมาตรฐานมีค่าเท่ากับ 1 ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป

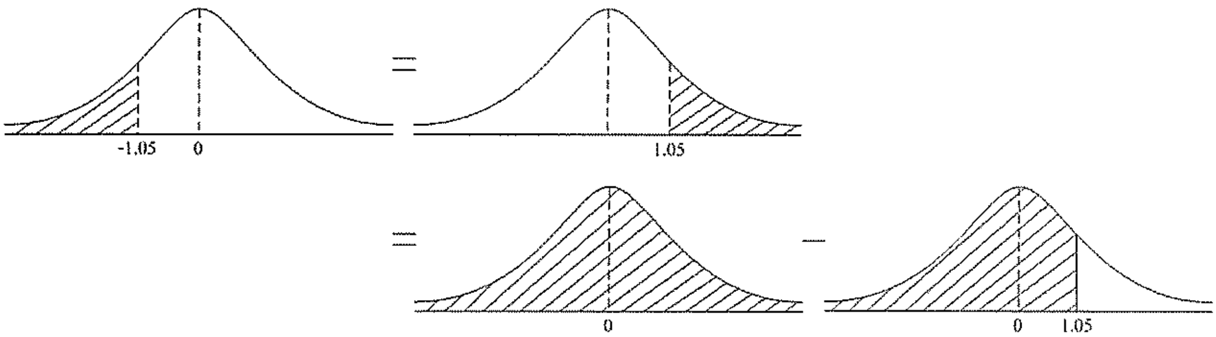


โดยใช้ตารางที่ 3.1 ได้ว่า

$$\begin{aligned} P(X \ge 1.65) &= 1 - P(X < 1.65) \\ &= 1 - 0.9505 \\ &= 0.0495 \end{aligned}$$

ดังนั้น $X \geq 1.65$ ด้วยความน่าจะเป็นเท่ากับ 0.0495

(3) พิจารณา $P(X < -1.05)$ จะเห็นว่ามีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X < -1.05$ เนื่องจากโค้งปกติมาตรฐานมีสมบัติสมมาตร และพื้นที่ใต้โค้งปกติมาตรฐานมีค่าเท่ากับ 1 ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป

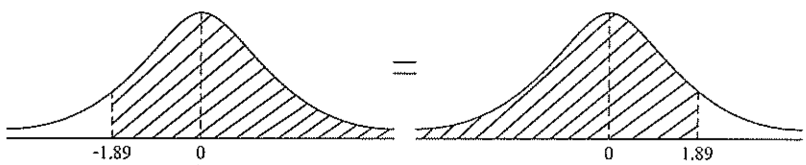


โดยใช้ตารางที่ 3.1 ได้ว่า

$$
\begin{aligned}
P(X < -1.05) &= P(X > 1.05) \\
&= 1 - P(X \leq 1.05) \\
&= 1 - 0.8531 \\
&= 0.1469
\end{aligned}
$$

ดังนั้น $X < -1.05$ ด้วยความน่าจะเป็นเท่ากับ 0.1469

(4) พิจารณา $P(X > -1.89)$ จะเห็นว่ามีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X > -1.89$ เนื่องจากโค้งปกติมาตรฐานมีสมบัติสมมาตร ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป

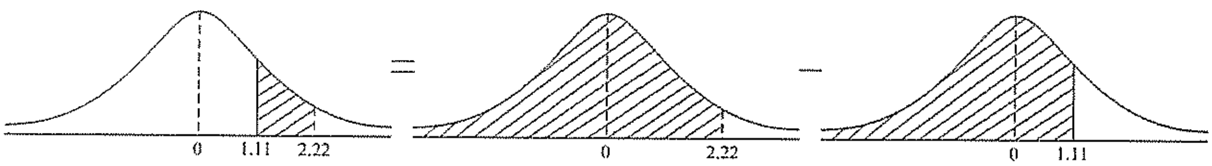


โดยใช้ตารางที่ 3.1 ได้ว่า

$$
\begin{aligned}
P(X > -1.89) &= P(X < 1.89) \\
&= 0.9706
\end{aligned}
$$

ดังนั้น $X > -1.89$ ด้วยความน่าจะเป็นเท่ากับ 0.9706

(5) พิจารณา $P(1.11 \leq X < 2.22)$ จะเห็นว่ามีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $1.11 \leq X < 2.22$ ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป

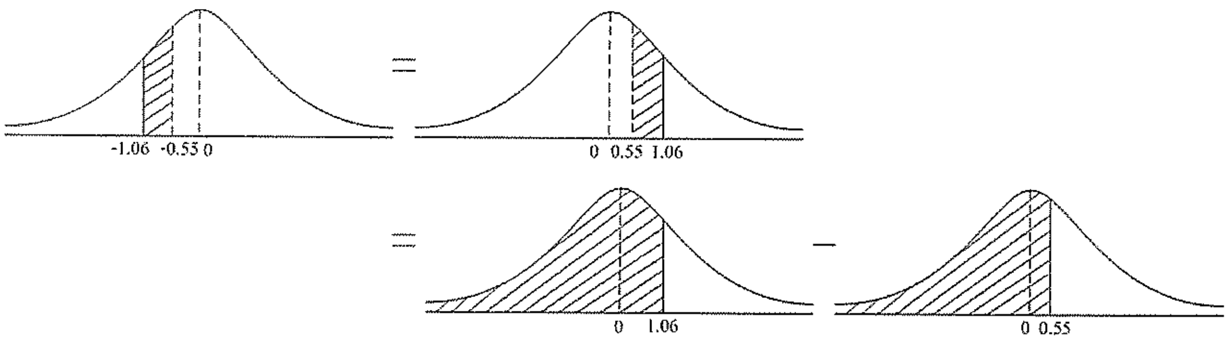


โดยใช้ตารางที่ 3.1 ได้ว่า

$$P(1.11 \le X < 2.22) = P(X < 2.22) - P(X < 1.11)$$
$$= 0.9868 - 0.8665$$
$$= 0.1203$$

ดังนั้น $1.11 \le X < 2.22$ ด้วยความน่าจะเป็นเท่ากับ 0.1203

(6) พิจารณา $P(-1.06 \le X < -0.55)$ จะเห็นว่ามีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $-1.06 \le X < -0.55$ เนื่องจากโค้งปกติมาตรฐานมีสมบัติสมมาตร ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป

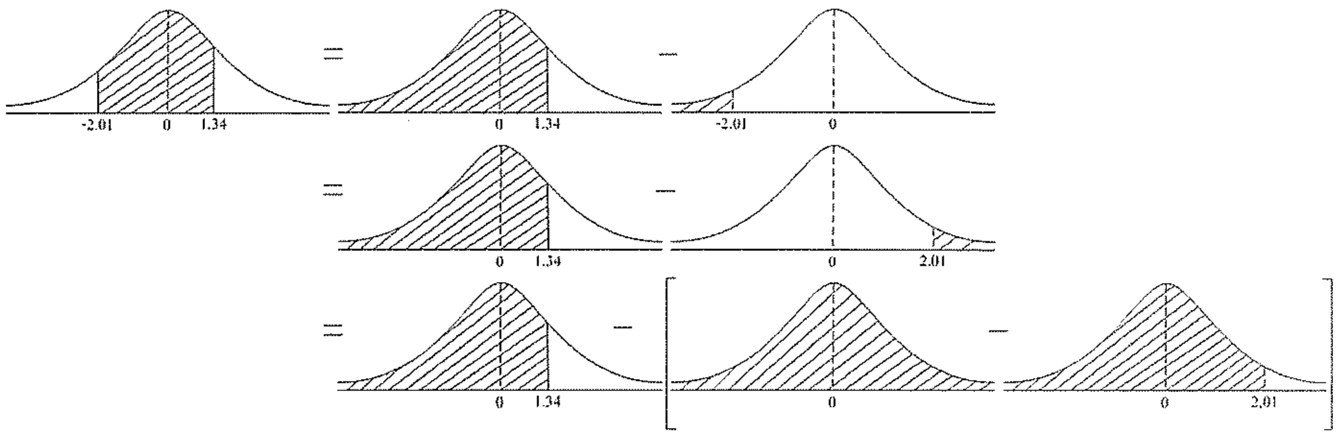


โดยใช้ตารางที่ 3.1 ได้ว่า

$$P(-1.06 \le X < -0.55) = P(0.55 < X \le 1.06)$$
$$= P(X \le 1.06) - P(X \le 0.55)$$
$$= 0.8554 - 0.7088$$
$$= 0.1466$$

ดังนั้น $-1.06 \le X < -0.55$ ด้วยความน่าจะเป็นเท่ากับ 0.1466

(7) พิจารณา $P(-2.01 \le X \le 1.34)$ จะเห็นว่ามีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐาน เนื่องจากโค้งปกติมาตรฐานมีสมบัติสมมาตร และพื้นที่ใต้โค้งปกติมาตรฐานมีค่าเท่ากับ 1 ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป



โดยใช้ตารางที่ 3.1 ได้ว่า

$$
\begin{aligned}
P(-2.01 \le X \le 1.34) &= P(X \le 1.34) - P(X < -2.01) \\
&= P(X \le 1.34) - P(X > 2.01) \\
&= P(X \le 1.34) - \{1 - P(X \le 2.01)\} \\
&= 0.9099 - \{1 - 0.9778\} \\
&= 0.8877
\end{aligned}
$$

ดังนั้น $-2.01 \le X \le 1.34$ ด้วยความน่าจะเป็นเท่ากับ 0.8877 ■

<u>ตัวอย่างที่ 4.2.2</u> กำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน นั่นคือ $X \sim N(0,1)$ จงหา

(1) ค่า a ที่ทำให้ $P(X \le a) = 0.8907$

(2) ค่า a ที่ทำให้ $P(X \ge a) = 0.1020$

(3) ค่า a ที่ทำให้ $P(X < a) = 0.1492$

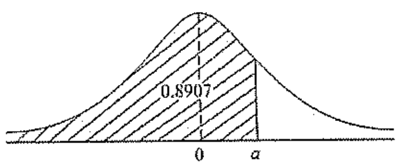
(4) ค่า a ที่ทำให้ $P(X > a) = 0.9783$

(5) ค่า a ที่ทำให้ $P(0 \le X < a) = 0.3264$

(6) ค่า a ที่ทำให้ $P(a < X < 0) = 0.4931$

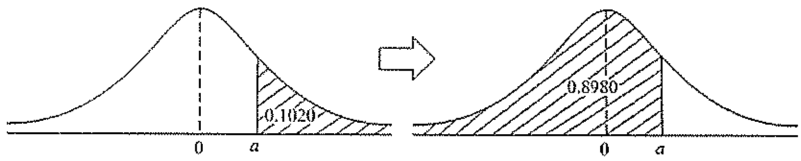
(7) ค่า a ที่ทำให้ $P(-a < X < a) = 0.4246$

วิธีทำ (1) พิจารณา $P(X \le a) = 0.8907$ หมายความว่า พื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X \le a$ มีค่าเท่ากับ 0.8907 แต่เนื่องจากพื้นที่ที่กำหนดให้มีค่ามากกว่า 0.5 ดังนั้น $a > 0$ ดังรูป



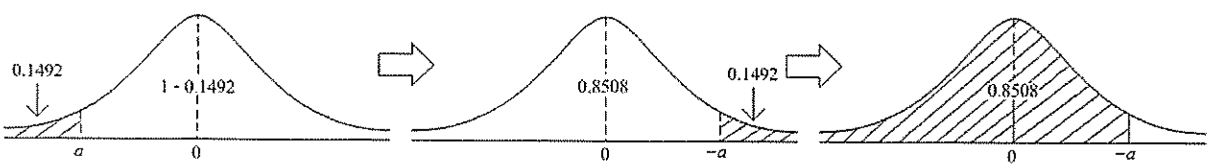
จากตารางที่ 3.1 จะได้ว่า $P(X \le 1.23) = 0.8907$ นั่นคือ $a = 1.23$
ดังนั้น ค่า a ที่ทำให้ $P(X \le a) = 0.8907$ มีค่าเท่ากับ 1.23

(2) พิจารณา $P(X \ge a) = 0.1020$ หมายความว่า พื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X \ge a$ มีค่าเท่ากับ 0.1020 แต่เนื่องจากพื้นที่ที่กำหนดให้มีค่าน้อยกว่า 0.5 ดังนั้น $a > 0$ และจากพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 1 จึงพิจารณาการหาค่า a ได้ดังรูป



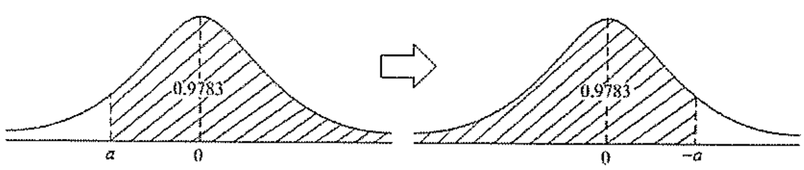
จากรูป $P(X < a) = 0.8980$ และจากตารางที่ 3.1 จะได้ว่า $P(X < 1.27) = 0.8980$ นั่นคือ $a = 1.27$
ดังนั้น ค่า a ที่ทำให้ $P(X \ge a) = 0.1020$ มีค่าเท่ากับ 1.27

(3) พิจารณา $P(X < a) = 0.1492$ หมายความว่า พื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X < a$ มีค่าเท่ากับ 0.1492 แต่เนื่องจากพื้นที่ที่กำหนดให้มีค่าน้อยกว่า 0.5 ดังนั้น $a < 0$ โดยสมบัติสมมาตร และจากพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 1 จึงพิจารณาการหาค่า a ได้ดังรูป



จากรูป $P(X \le -a) = 0.8508$ และจากตารางที่ 3.1 จะได้ว่า $P(X \le 1.04) = 0.8508$ จึงได้ว่า $-a = 1.04$ นั่นคือ $a = 1.04$
ดังนั้น ค่า a ที่ทำให้ $P(X < a) = 0.1492$ มีค่าเท่ากับ 1.04

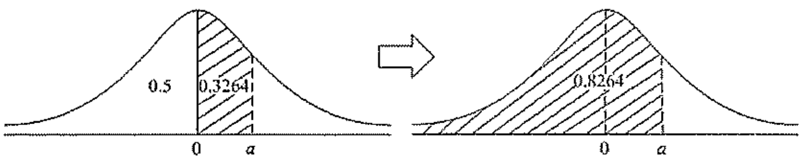
(4) พิจารณา $P(X > a) = 0.9783$ หมายความว่า พื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $X > a$ มีค่าเท่ากับ 0.9783 แต่เนื่องจากพื้นที่ที่กำหนดให้มีค่ามากกว่า 0.5 ดังนั้น $a < 0$ โดยสมบัติสมมาตร จึงพิจารณาการหาค่า a ได้ดังรูป



จากรูป $P(X<-a)=0.9783$ และจากตารางที่ 3.1 จะได้ว่า $P(X<2.02)=0.9783$ จึงได้ว่า $-a=2.02$ นั่นคือ $a=2.02$

ดังนั้น ค่า a ที่ทำให้ $P(X>a)=0.9783$ มีค่าเท่ากับ 2.02

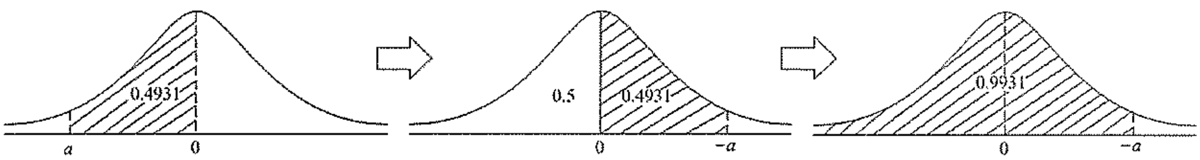
(5) พิจารณา $P(0\le X<a)=0.3264$ หมายความว่า พื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $0\le X<a$ มีค่าเท่ากับ 0.3264 โดยสมบัติสมมาตร และจากพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 1 (ครึ่งหนึ่งของพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 0.5) จึงพิจารณาการหาค่า a ได้ดังรูป



จากรูป $P(X<a)=0.8264$ และจากตารางที่ 3.1 จะได้ว่า $P(X<0.94)=0.8264$ นั่นคือ $a=0.94$

ดังนั้น ค่า a ที่ทำให้ $P(0\le X<a)=0.3264$ มีค่าเท่ากับ 0.94

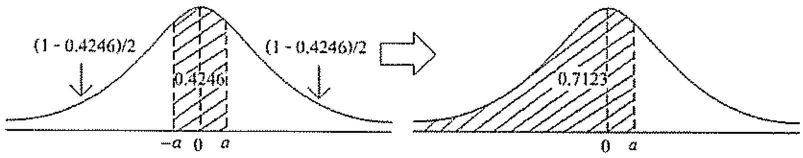
(6) พิจารณา $P(a<X<0)=0.4931$ หมายความว่า พื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $a<X<0$ มีค่าเท่ากับ 0.4931 โดยสมบัติสมมาตร และจากพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 1 (ครึ่งหนึ่งของพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 0.5) จึงพิจารณาการหาค่า a ได้ดังรูป



จากรูป $P(X<-a)=0.9931$ และจากตารางที่ 3.1 จะได้ว่า $P(X<2.46)=0.9931$ จึงได้ว่า $-a=2.46$ นั่นคือ $a=2.46$

ดังนั้น ค่า a ที่ทำให้ $P(a<X<0)=0.4931$ มีค่าเท่ากับ 2.46

(7) พิจารณา $P(-a<X<a)=0.4246$ หมายความว่า พื้นที่ใต้โค้งปกติมาตรฐาน โดยที่ $-a<X<a$ มีค่าเท่ากับ 0.4246 โดยสมบัติสมมาตร และจากพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 1 (ครึ่งหนึ่งของพื้นที่ใต้โค้งปกติมีค่าเท่ากับ 0.5) จึงพิจารณาการหาค่า a ได้ดังรูป



จากรูป $P(X<a)=0.7123$ และจากตารางที่ 3.1 จะได้ว่า $P(X<0.56)=0.7123$ นั่นคือ $a=0.56$

ดังนั้น ค่า a ที่ทำให้ $P(-a<X<a)=0.4246$ มีค่าเท่ากับ 0.56 ■

ตัวอย่างที่ 4.2.3 กำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติโดยมีค่าเฉลี่ยเท่ากับ 1 และความแปรปรวนเท่ากับ 4 นั่นคือ $X\sim N(1,4)$ จงหา

(1) $P(X>0)$

(2) $P(-0.23<X<1.56)$

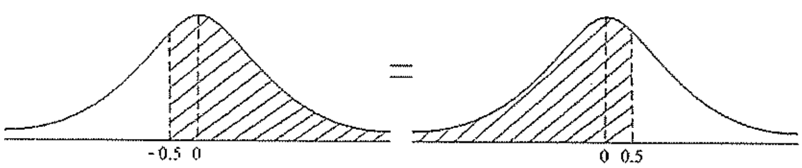
วิธีทำ เนื่องจากตัวแปรสุ่ม X มีค่าเฉลี่ยเท่ากับ 1 และความแปรปรวนเท่ากับ 4 นั่นคือ $\mu=1$ และ $\sigma^2=4$ เพื่อให้สามารถใช้ตารางที่ 3.1 ช่วยในการหาค่าความน่าจะเป็น ดังนั้นเราจะแปลงตัวแปรสุ่ม X ไปเป็นตัวแปรสุ่ม Z ซึ่งมีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐานได้โดยใช้สมการ

$$Z=\frac{X-\mu}{\sigma}$$

(1) แปลงจาก $X\sim N(1,4)$ ไปเป็น $Z\sim N(0,1)$ ได้ดังนี้

$$P(X>0)=P(\frac{X-1}{2}>\frac{0-1}{2})=P(Z>-0.5)$$

พิจารณา $P(Z>-0.5)$ มีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐานโดยที่ $Z>-0.5$ เนื่องจากโค้งปกติมาตรฐานมีสมบัติสมมาตร ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป



โดยใช้ตารางที่ 3.1 ได้ว่า

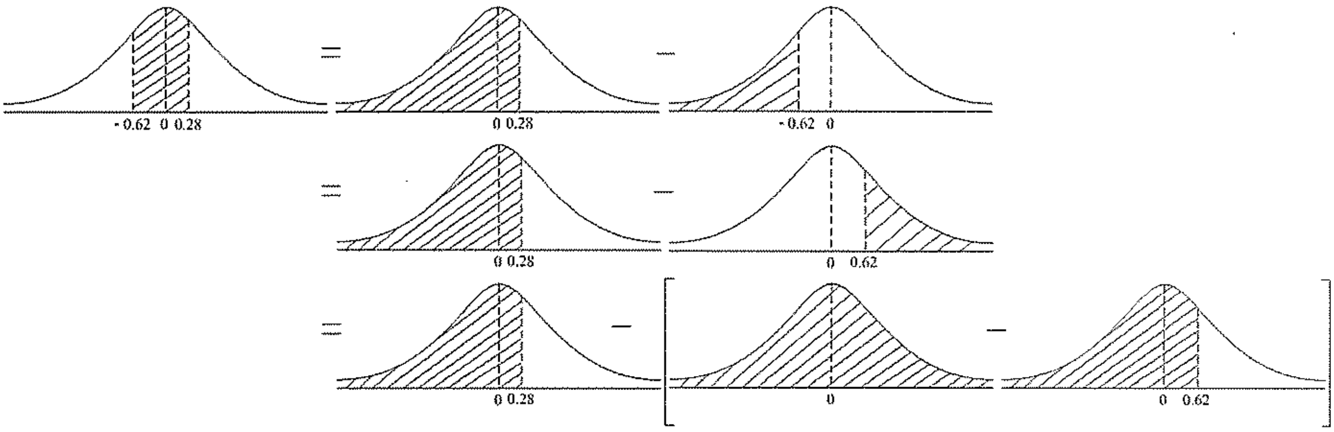
$$\begin{aligned}P(Z>-0.5)&=P(Z<0.5)\\&=0.6915\end{aligned}$$

ดังนั้น $X>0$ ด้วยความน่าจะเป็นเท่ากับ 0.6915

(2) แปลงจาก $X\sim N(1,4)$ ไปเป็น $Z\sim N(0,1)$ ได้ดังนี้

$$P(-0.23 < X < 1.56) = P(\frac{-0.23-1}{2} < \frac{X-1}{2} < \frac{1.56-1}{2}) = P(-0.62 < Z < 0.28)$$

พิจารณา $P(-0.62 < Z < 0.28)$ มีค่าเท่ากับพื้นที่ใต้โค้งปกติมาตรฐานโดยที่ $-0.62 < Z < 0.28$ เนื่องจากโค้งปกติมาตรฐานมีสมบัติสมมาตร และพื้นที่ใต้โค้งปกติมาตรฐานมีค่าเท่ากับ 1 ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป



โดยใช้ตารางที่ 3.1 ได้ว่า

$$\begin{aligned} P(-0.62 < Z < 0.28) &= P(Z < 0.28) - P(Z \le -0.62) \\ &= P(Z < 0.28) - P(Z \ge 0.62) \\ &= P(Z < 0.28) - \{1 - P(Z < 0.62)\} \\ &= 0.6103 - \{1 - 0.7324\} \\ &= 0.3427 \end{aligned}$$

ดังนั้น $-0.23 < X < 1.56$ ด้วยความน่าจะเป็นเท่ากับ 0.3427 ■

ตัวอย่างที่ 4.2.4 กำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติโดยมีค่าเฉลี่ยเท่ากับ 1 และความแปรปรวนเท่ากับ 4 นั่นคือ $X \sim N(1,4)$ จงหาค่า a ที่ทำให้ $P(X > a) = 0.7123$

วิธีทำ พิจารณาการหาค่า a ที่ทำให้

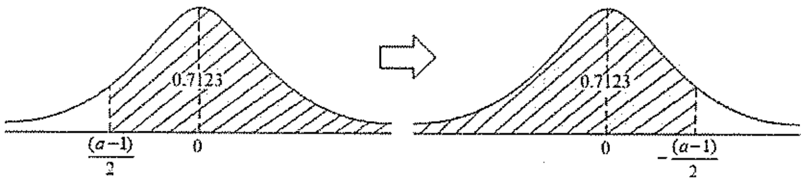
$$P(X > a) = 0.7123$$

เพื่อให้สามารถใช้ตารางที่ 3.1 ช่วยในการหาค่า a ดังนั้นเราจะแปลงตัวแปรสุ่ม X ไปเป็นตัวแปรสุ่ม Z ซึ่งมีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐานได้ดังนี้

$$P(\frac{X-1}{2} > \frac{a-1}{2}) = 0.7123$$

$$P(Z > \frac{a-1}{2}) = 0.7123$$

แต่เนื่องจากพื้นที่ที่กำหนดให้มีค่ามากกว่า 0.5 ดังนั้น $\frac{a-1}{2}<0$ โดยสมบัติสมมาตร จึงพิจารณาการหาค่า a ได้ดังรูป



จากรูป $P(Z<-\frac{a-1}{2})=0.7123$ และจากตารางที่ 3.1 จะได้ว่า $P(Z<0.56)=0.7123$ จึงได้ว่า $-\frac{a-1}{2}=0.56$ นั่นคือ $a=-0.12$

ดังนั้น ค่า a ที่ทำให้ $P(X>a)=0.7123$ มีค่าเท่ากับ – 0.12 ■

<u>ตัวอย่างที่ 4.2.5</u> เครื่องจักรของโรงงานแห่งหนึ่งมีอายุการใช้งานโดยเฉลี่ย 5 ปี ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ปี ถ้าอายุการใช้งานของเครื่องจักรมีการแจกแจงความน่าจะเป็นแบบปกติ จงหา

(1) ความน่าจะเป็นที่เครื่องจักรนี้จะมีอายุการใช้งานมากกว่า 3 ปี

(2) ความน่าจะเป็นที่เครื่องจักรนี้จะมีอายุการใช้งานน้อยกว่า 4 ปี

(3) ความน่าจะเป็นที่เครื่องจักรนี้จะมีอายุการใช้งานตั้งแต่ 2 - 4 ปี

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนอายุการใช้งานของเครื่องจักร ดังนั้น

$$X \sim N(5,1)$$

(1) ความน่าจะเป็นที่เครื่องจักรนี้จะมีอายุการใช้งานมากกว่า 3 ปี เขียนแทนได้ด้วยสัญลักษณ์ $P(X>3)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X>3) &= P(\frac{X-5}{1}>\frac{3-5}{1}) \\ &= P(Z>-2) \\ &= P(Z<2) \\ &= 0.9972 \end{aligned}$$

ดังนั้น เครื่องจักรนี้จะมีอายุการใช้งานมากกว่า 3 ปี ด้วยความน่าจะเป็นเท่ากับ 0.9972

(2) ความน่าจะเป็นที่เครื่องจักรนี้จะมีอายุการใช้งานน้อยกว่า 4 ปี เขียนแทนได้ด้วยสัญลักษณ์ $P(X<4)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X<4) &= P(\frac{X-5}{1}<\frac{4-5}{1}) \\ &= P(Z<-1) \end{aligned}$$

$$= P(Z > 1)$$
$$= 1 - P(Z < 1)$$
$$= 1 - 0.8413$$
$$= 0.1587$$

ดังนั้น เครื่องจักรนี้จะมีอายุการใช้งานน้อยกว่า 4 ปี ด้วยความน่าจะเป็นเท่ากับ 0.1587

(3) ความน่าจะเป็นที่เครื่องจักรนี้จะมีอายุการใช้งานตั้งแต่ 2 - 4 ปี เขียนแทนได้ด้วยสัญลักษณ์ $P(2 \le X \le 4)$ ซึ่งหาค่าได้ดังนี้

$$P(2 \le X \le 4) = P(\frac{2-5}{1} \le \frac{X-5}{1} \le \frac{4-5}{1})$$
$$= P(-3 \le Z \le -1)$$
$$= P(1 \le Z \le 3)$$
$$= P(Z \le 3) - P(Z < 1)$$
$$= 0.9987 - 0.8413$$
$$= 0.1574$$

ดังนั้น เครื่องจักรนี้จะมีอายุการใช้งานตั้งแต่ 2 - 4 ปี ด้วยความน่าจะเป็นเท่ากับ 0.1574 ■

ตัวอย่างที่ 4.2.6 โรงงานผลิตน้ำอัดลมแห่งหนึ่ง ใช้เครื่องจักรในการบรรจุน้ำอัดลมพบว่า ปริมาณน้ำอัดลมมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ย 200 มิลลิลิตรต่อขวด ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 15 มิลลิลิตรต่อขวด จงหา

(1) ความน่าจะเป็นที่น้ำอัดลมที่ผลิตจากโรงงานแห่งนี้จะมีปริมาณน้อยกว่า 176 มิลลิลิตรต่อขวด

(2) ความน่าจะเป็นที่น้ำอัดลมที่ผลิตจากโรงงานแห่งนี้จะมีปริมาณมากกว่า 180 มิลลิลิตรต่อขวด

(3) ความน่าจะเป็นที่น้ำอัดลมที่ผลิตจากโรงงานแห่งนี้จะมีปริมาณตั้งแต่ 190 – 212 มิลลิลิตรต่อขวด

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนปริมาณน้ำอัดลมแต่ละขวดที่ผลิตจากโรงงานแห่งนี้ ดังนั้น

$$X \sim N(200, (15)^2)$$

(1) ความน่าจะเป็นที่น้ำอัดลมจะมีปริมาณมากกว่า 176 มิลลิลิตรต่อขวด เขียนแทนได้ด้วยสัญลักษณ์ $P(X < 176)$ ซึ่งหาค่าได้ดังนี้

$$P(X < 176) = P(\frac{X-200}{15} < \frac{176-200}{15})$$
$$= P(Z < -1.6)$$
$$= P(Z > 1.6)$$
$$= 1 - P(Z \le 1.6)$$
$$= 1 - 0.9452$$
$$= 0.0548$$

ดังนั้น น้ำอัดลมที่ผลิตจากโรงงานแห่งนี้จะมีปริมาณน้อยกว่า 176 มิลลิลิตรต่อขวด ด้วยความน่าจะเป็นเท่ากับ 0.0548

(2) ความน่าจะเป็นที่น้ำอัดลมจะมีปริมาณมากกว่า 180 มิลลิลิตรต่อขวด เขียนแทนได้ด้วยสัญลักษณ์ $P(X>180)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X>180) &= P(\frac{X-200}{15} > \frac{180-200}{15}) \\ &= P(Z>-1.33) \\ &= P(Z<1.33) \\ &= 0.9082 \end{aligned}$$

ดังนั้น น้ำอัดลมที่ผลิตจากโรงงานแห่งนี้จะมีปริมาณมากกว่า 180 มิลลิลิตรต่อขวด ด้วยความน่าจะเป็นเท่ากับ 0.9082

(3) ความน่าจะเป็นที่น้ำอัดลมจะมีปริมาณตั้งแต่ 190 – 212 มิลลิลิตรต่อขวด เขียนแทนได้ด้วยสัญลักษณ์ $P(190 \le X \le 212)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(190 \le X \le 212) &= P(\frac{190-200}{15} \le \frac{X-200}{15} \le \frac{212-200}{15}) \\ &= P(-0.67 \le Z \le 0.8) \\ &= P(Z \le 0.8) - P(Z < -0.67) \\ &= P(Z \le 0.8) - P(Z > 0.67) \\ &= P(Z \le 0.8) - \{1 - P(Z \le 0.67)\} \\ &= 0.7881 - \{1 - 0.7486\} \\ &= 0.5367 \end{aligned}$$

ดังนั้น เครื่องจักรนี้จะมีอายุการใช้งานตั้งแต่ 2 - 4 ปี ด้วยความน่าจะเป็นเท่ากับ 0.5367 ■

<u>ตัวอย่างที่ 4.2.7</u> ระยะเวลาที่ใช้ในการเดินทางจากบ้านไปยังโรงเรียนของเด็กคนหนึ่งมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ย 24 นาที ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 3.8 นาที จงหา

(1) ความน่าจะเป็นที่เด็กคนนี้จะใช้เวลาในการเดินทางจากบ้านไปยังโรงเรียนน้อยกว่าครึ่งชั่วโมง

(2) ถ้าเวลาเริ่มเรียนคือ 09.00 น. และเขาเริ่มออกจากบ้านเวลา 08.45 น. จงหาความน่าจะเป็นที่เด็กคนนี้จะมาเรียนสาย

(3) ถ้าคุณครูนำขนมมาแจกที่หน้าโรงเรียนตั้งแต่เวลา 08.50 น. – 09.00 น. และเขาเริ่มออกจากบ้านเวลา 08.35 น. จงหาความน่าจะเป็นที่เด็กคนนี้จะไม่ได้รับขนม

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนระยะเวลาที่ใช้ในการเดินทางจากบ้านไปยังที่ทำงานของชายคนนี้ ดังนั้น

$$X \sim N(24, (3.8)^2)$$

(1) ความน่าจะเป็นที่เด็กคนนี้จะใช้เวลาในการเดินทางจากบ้านไปยังโรงเรียนน้อยกว่าครึ่งชั่วโมง เขียนแทนได้ด้วยสัญลักษณ์ $P(X<30)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X<30) &= P(\frac{X-24}{3.8}<\frac{30-24}{3.8}) \\ &= P(Z<1.58) \\ &= 0.9429 \end{aligned}$$

ดังนั้น เด็กคนนี้จะใช้เวลาในการเดินทางจากบ้านไปยังโรงเรียนน้อยกว่าครึ่งชั่วโมงด้วยความน่าจะเป็นเท่ากับ 0.9429

(2) ถ้าเวลาเริ่มเรียนคือ 09.00 น. และเขาเริ่มออกจากบ้านเวลา 08.45 น. นั่นคือ เขาจะมาเรียนสายถ้าใช้เวลาเดินทางมากกว่า 15 นาที ดังนั้น ความน่าจะเป็นที่เด็กคนนี้จะมาเรียนสาย เขียนแทนได้ด้วยสัญลักษณ์ $P(X>15)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(X>15) &= P(\frac{X-24}{3.8}>\frac{15-24}{3.8}) \\ &= P(Z>-2.37) \\ &= P(Z<2.37) \\ &= 0.9911 \end{aligned}$$

ดังนั้น เด็กคนนี้จะมาเรียนสายด้วยความน่าจะเป็นเท่ากับ 0.9911

(3) ถ้าคุณครูนำขนมมาแจกที่หน้าโรงเรียนตั้งแต่เวลา 08.50 น. – 09.00 น. และเขาเริ่มออกจากบ้านเวลา 08.35 น. นั่นคือ เขาจะต้องใช้เวลาเดินทางมาโรงเรียนตั้งแต่ 15 -25 นาที จึงจะได้รับขนม พิจารณาความน่าจะเป็นที่เด็กคนนี้จะได้รับขนม เขียนแทนได้ด้วยสัญลักษณ์ $P(15\le X\le 25)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(15\le X\le 25) &= P(\frac{15-24}{3.8}\le\frac{X-24}{3.8}\le\frac{25-24}{3.8}) \\ &= P(-2.37\le Z\le 0.26) \\ &= P(Z\le 0.26)-P(Z<-2.37) \\ &= P(Z\le 0.26)-P(Z>2.37) \\ &= P(Z\le 0.26)-\{1-P(Z\le 2.37)\} \\ &= 0.6026-\{1-0.9911\} \\ &= 0.5937 \end{aligned}$$

ดังนั้น เด็กคนนี้จะไม่ได้รับขนมด้วยความน่าจะเป็นเท่ากับ $1-0.5937=0.4063$ ■

__ตัวอย่างที่ 4.2.8__ ถ้าคะแนนสอบวิชาความน่าจะเป็นและสถิติมีการแจกแจงความน่าจะเป็นแบบปกติด้วยคะแนนเฉลี่ย 60 คะแนน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 15 คะแนน จงหา

(1) นักศึกษาที่สอบได้คะแนนตั้งแต่ 85 - 95 คะแนน มีทั้งหมดกี่เปอร์เซ็นต์

(2) ถ้าเกณฑ์การตัดเกรดของวิชานี้คือ นักศึกษาที่สอบได้คะแนนต่ำสุด 10% ของนักศึกษาทั้งหมด จะได้เกรด F อยากทราบว่านักศึกษาจะต้องได้คะแนนมากกว่าเท่าใดจึงจะไม่ได้เกรด F

(3) ถ้าเกณฑ์การตัดเกรดของวิชานี้คือ นักศึกษาที่สอบได้คะแนนสูงสุด 5% ของนักศึกษาทั้งหมด จะได้เกรด A อยากทราบว่านักศึกษาจะต้องได้คะแนนอย่างน้อยเท่าใดจึงจะได้เกรด A

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนคะแนนสอบวิชาความน่าจะเป็นและสถิติ ดังนั้น

$$X \sim N(60,(15)^2)$$

(1) ความน่าจะเป็นที่นักศึกษาจะสอบได้คะแนนตั้งแต่ 85 - 95 คะแนน แทนได้ด้วยสัญลักษณ์ $P(85 \le X \le 95)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(85 \le X \le 95) &= P\left(\frac{85-60}{15} \le \frac{X-60}{15} \le \frac{95-60}{15}\right) \\ &= P(1.67 \le Z \le 2.33) \\ &= P(Z \le 2.33) - P(Z < 1.67) \\ &= 0.9901 - 0.9525 = 0.0376 \end{aligned}$$

ดังนั้น นักศึกษาที่สอบได้คะแนนตั้งแต่ 85 - 95 คะแนน คิดเป็น 3.76% ของนักศึกษาทั้งหมด

(2) กำหนดให้ a แทนคะแนนที่น้อยที่สุดที่จะไม่ได้เกรด F

เนื่องจากนักศึกษาที่สอบได้คะแนนต่ำสุด 10% ของนักศึกษาทั้งหมด จะได้เกรด F หมายความว่าต้องการหาค่า a ที่ทำให้

$$P(X \le a) = 0.1$$

หรือ

$$P(Z \le \frac{a-60}{15}) = 0.1$$

เนื่องจากโค้งปกติมาตรฐานมีสมบัติสมมาตรจึงได้ว่า

$$P(Z \ge -\frac{a-60}{15}) = 0.1$$

และจากพื้นที่ใต้โค้งปกติมาตรฐานมีค่าเท่ากับ 1 จึงได้ว่า

$$P(Z < -\frac{a-60}{15}) = 0.9$$

จากตารางที่ 3.2 จะพบว่า

$$P(Z < 1.282) = 0.9$$

ดังนั้น

$$-\frac{a-60}{15} = 1.282$$

นั่นคือ

$$a = -(15)(1.282) + 60 = 40.77$$

เพราะฉะนั้น นักศึกษาจะต้องได้คะแนนมากกว่า 40.77 คะแนน จึงจะไม่ได้เกรด F

(3) กำหนดให้ a แทนคะแนนที่น้อยที่สุดที่จะได้เกรด A

เนื่องจากนักศึกษาที่สอบได้คะแนนสูงสุด 5% ของนักศึกษาทั้งหมด จะได้เกรด A หมายความว่าต้องการหาค่า a ที่ทำให้

$$P(X \geq a) = 0.05$$

หรือ

$$P(Z \geq \frac{a-60}{15}) = 0.05$$

เนื่องจากพื้นที่ใต้โค้งปกติมาตรฐานมีค่าเท่ากับ 1 จึงได้ว่า

$$P(Z < \frac{a-60}{15}) = 0.95$$

จากตารางที่ 3.2 จะพบว่า

$$P(Z < 1.645) = 0.95$$

ดังนั้น

$$\frac{a-60}{15} = 1.645$$

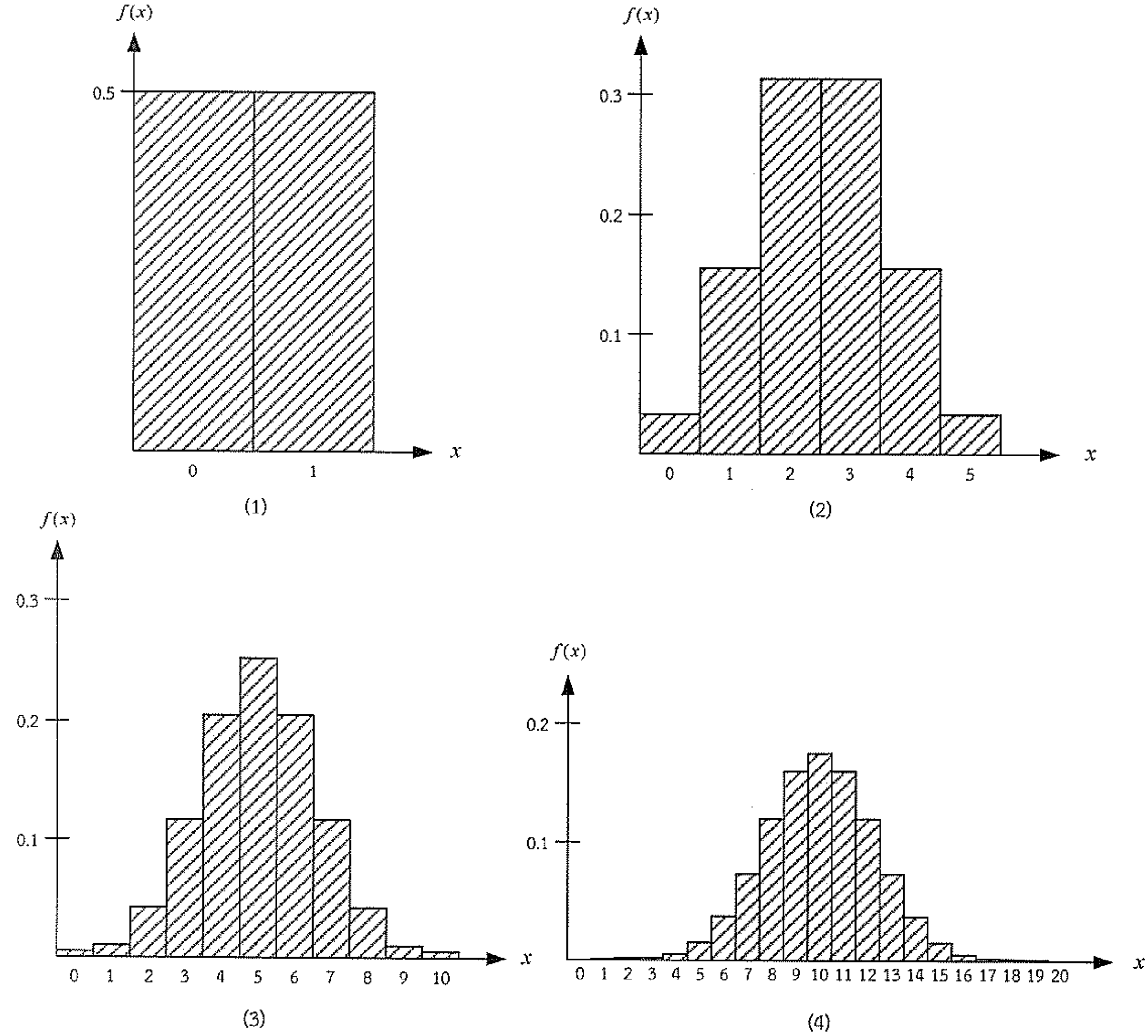
นั่นคือ

$$a = (15)(1.645) + 60 = 84.675$$

เพราะฉะนั้น นักศึกษาจะต้องได้คะแนนอย่างน้อย 84.675 คะแนน จึงจะได้เกรด A ■

4.2.1 การประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปกติ

กรณีที่ $X \sim B(n, p)$ และ n มีค่ามาก p มีค่าน้อยเข้าใกล้ศูนย์ เราสามารถประมาณค่าความน่าจะเป็นแบบทวินามได้ด้วยความน่าจะเป็นแบบปัวซง ตามที่ได้กล่าวถึงแล้วในบทที่ 3 แต่สำหรับในหัวข้อนี้จะกล่าวถึงกรณีที่ $X \sim B(n, p)$ และ n มีค่ามาก แต่ p มีค่าไม่เข้าใกล้ศูนย์แต่เข้าใกล้ 0.5 พิจารณารูปที่ 4.2.7



รูปที่ 4.2.7 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทวินาม นั่นคือ $X \sim B(n, 0.5)$ เมื่อ (1) $n=1$ (2) $n=5$ (3) $n=10$ (4) $n=20$

จากรูปที่ 4.2.7. จะสังเกตว่า เมื่อ n มีค่ามากขึ้น และ $p=0.5$ ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทวินามจะใกล้เคียงกับฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติ

กรณีที่ $X \sim B(n, p)$ และ n มีค่ามาก และ p มีค่าไม่เข้าใกล้ศูนย์หรือหนึ่งแต่กลับมีค่าใกล้เคียงกับ 0.5 ดังนั้นเราสามารถประมาณความน่าจะเป็นแบบทวินามได้ด้วยความน่าจะเป็นแบบปกติ โดยที่ $\mu = np$ และ $\sigma^2 = np(1-p)$

เนื่องจากการใช้ความน่าจะเป็นแบบปกติเพื่อประมาณความน่าจะเป็นแบบทวินาม เป็นการใช้ตัวแปรสุ่มแบบต่อเนื่องมาประมาณตัวแปรสุ่มไม่ต่อเนื่อง ดังนั้น การประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะ

เป็นแบบปกติจึงต้องมีการปรับค่าเล็กน้อย เพื่อให้ได้ค่าที่ใกล้เคียงมากขึ้น ค่าที่ใช้ในการปรับ คือ 0.5 ซึ่งการปรับค่าดังกล่าวเรียกว่า continuity correction โดยสามารถปรับค่าได้ตามตารางที่ 4.2.1 ดังนี้

ทวินาม	ปกติ
$P(a<X<b)$	$P(a+0.5\le X\le b-0.5)$
$P(a\le X\le b)$	$P(a-0.5\le X\le b+0.5)$
$P(a\le X<b)$	$P(a-0.5\le X\le b-0.5)$
$P(a<X\le b)$	$P(a+0.5\le X\le b+0.5)$
$P(X>a)$	$P(X\ge a+0.5)$
$P(X\ge a)$	$P(X\ge a-0.5)$
$P(X<a)$	$P(X\le a-0.5)$
$P(X\le a)$	$P(X\le a+0.5)$

ตารางที่ 4.2.1 การปรับค่าด้วย 0.5 เมื่อใช้ความน่าจะเป็นแบบปกติมาประมาณความน่าจะเป็นแบบทวินาม

ตัวอย่างที่ 4.2.9 กำหนดให้ตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบทวินาม โดยที่ $n=100$ และ $p=\frac{1}{2}$ นั่นคือ $X\sim B(100,\frac{1}{2})$ จงหา

(1) $P(40\le X\le 59)$

(2) $P(52\le X<55)$

(3) $P(42<X\le 45)$

(4) $P(49<X<51)$

(5) $P(X\le 38)$

(6) $P(X<63)$

(7) $P(X>47)$

(8) $P(X\ge 36)$

วิธีทำ ในตัวอย่างนี้จะแสดงวิธีหาค่าความน่าจะเป็น 3 วิธี ดังนี้

วิธีที่ 1 คำนวณโดยตรงจากฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม

เนื่องจาก $X\sim B(100,\frac{1}{2})$ ดังนั้น ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการ

$$f(x,100,0.5)=\binom{100}{x}(0.5)^x(1-0.5)^{100-x}$$

(1) $P(40\le X\le 59)=P(X=40)+\ldots+P(X=59)$

(6) $$\begin{aligned} P(X<63) &= P(X \le 62.5) \\ &= P\left(\frac{X-50}{5} \le \frac{62.5-50}{5}\right) \\ &= P(Z \le 2.5) \\ &= 0.9938 \end{aligned}$$

ดังนั้น $X<63$ ด้วยความน่าจะเป็นเท่ากับ 0.9938

(7) $$\begin{aligned} P(X>47) &= P(X \ge 47.5) \\ &= P\left(\frac{X-50}{5} \ge \frac{47.5-50}{5}\right) \\ &= P(Z \ge -0.5) \\ &= 0.6915 \end{aligned}$$

ดังนั้น $X>47$ ด้วยความน่าจะเป็นเท่ากับ 0.6915

(8) $$\begin{aligned} P(X \ge 36) &= P(X \ge 35.5) \\ &= P\left(\frac{X-50}{5} \ge \frac{35.5-50}{5}\right) \\ &= P(Z \ge -2.9) \\ &= 0.9981 \end{aligned}$$

ดังนั้น $X \ge 36$ ด้วยความน่าจะเป็นเท่ากับ 0.9981 ■

<u>ตัวอย่างที่ 4.2.10</u> จากการศึกษาข้อมูลของนักศึกษาในมหาวิทยาลัยแห่งหนึ่งพบว่า 43% ของนักศึกษาจะจบหลักสูตรภายใน 4 ปี ถ้ามีการสุ่มนักศึกษามาจำนวน 200 คน เพื่อติดตามผล จงหา

(1) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ตั้งแต่ 40 ถึง 100 คน

(2) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ระหว่าง 75 กับ 95 คน

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนจำนวนนักศึกษาที่จบการศึกษาภายใน 4 ปี ดังนั้น

$$X \sim B(200, 0.43)$$

ในตัวอย่างนี้จะแสดงวิธีหาค่าความน่าจะเป็น 3 วิธี ดังนี้

<u>วิธีที่ 1</u> คำนวณโดยตรงจากฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม

เนื่องจาก $X \sim B(200, 0.43)$ ดังนั้น ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการ

$$f(x, 200, 0.43) = \binom{200}{x}(0.43)^x(1-0.43)^{200-x}$$

(1) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ตั้งแต่ 40 ถึง 100 คน แทนได้ด้วยสัญลักษณ์ $P(40 \le X \le 100)$ ซึ่งหาค่าได้ดังนี้

$$P(40 \le X \le 100) = P(X=40) + \ldots + P(X=100)$$

$$= \binom{200}{40}(0.43)^{40}(1-0.43)^{200-40} + \ldots + \binom{200}{100}(0.43)^{100}(1-0.43)^{200-100}$$

$$= 0.9804$$

ดังนั้น จะนักศึกษาจบที่ภายใน 4 ปี ตั้งแต่ 40 ถึง 100 คน ด้วยความน่าจะเป็นเท่ากับ 0.9804

(2) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ระหว่าง 75 กับ 95 คน แทนได้ด้วยสัญลักษณ์ $P(75 < X < 95)$ ซึ่งหาค่าได้ดังนี้

$$P(75 < X < 95) = P(X = 76) + \ldots + P(X = 94)$$

$$= \binom{200}{76}(0.43)^{76}(1-0.43)^{200-76} + \ldots + \binom{200}{94}(0.43)^{94}(1-0.43)^{200-94}$$

$$= 0.8212$$

ดังนั้น จะมีนักศึกษาที่จบภายใน 4 ปี ระหว่าง 75 กับ 95 คน ด้วยความน่าจะเป็นเท่ากับ 0.8212

วิธีที่ 2 คำนวณโดยใช้การประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปกติโดยไม่มีการปรับค่าด้วย 0.5

เนื่องจาก $X \sim B(200, 0.43)$ โดยที่ $n = 200$ มีค่ามาก และ $p = 0.43$ มีค่าไม่เข้าใกล้ศูนย์ ดังนั้น เราสามารถประมาณค่าความน่าจะเป็นแบบทวินามได้ด้วยความน่าจะเป็นแบบปกติ โดยที่ $\mu = np = (200)(0.43) = 86$ และ $\sigma^2 = np(1-p) = (200)(0.43)(1-0.43) = 49.02$ นั่นคือ $\sigma = \sqrt{49.02}$

(1) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ตั้งแต่ 40 ถึง 100 คน แทนได้ด้วยสัญลักษณ์ $P(40 \le X \le 100)$ ซึ่งหาค่าได้ดังนี้

$$P(40 \le X \le 100) = P\left(\frac{40-86}{\sqrt{49.02}} \le X \le \frac{100-86}{\sqrt{49.02}}\right)$$

$$= P(-6.57 \le Z \le 2)$$

$$= 0.9772$$

ดังนั้น จะนักศึกษาจบที่ภายใน 4 ปี ตั้งแต่ 40 ถึง 100 คน ด้วยความน่าจะเป็นเท่ากับ 0.9772

(2) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ระหว่าง 75 กับ 95 คน แทนได้ด้วยสัญลักษณ์ $P(75 < X < 95)$ ซึ่งหาค่าได้ดังนี้

$$P(75 < X < 95) = P\left(\frac{75-86}{\sqrt{49.02}} \le X \le \frac{95-86}{\sqrt{49.02}}\right)$$

$$= P(-1.57 \le Z \le 1.29)$$

$$= 0.8433$$

ดังนั้น จะมีนักศึกษาที่จบภายใน 4 ปี ระหว่าง 75 กับ 95 คน ด้วยความน่าจะเป็นเท่ากับ 0.8433

วิธีที่ 3 คำนวณโดยใช้การประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปกติโดยมีการปรับค่าด้วย 0.5

เนื่องจาก $X \sim B(200, 0.43)$ โดยที่ $n = 200$ มีค่ามาก และ $p = 0.43$ มีค่าไม่เข้าใกล้ศูนย์ ดังนั้น เราสามารถประมาณค่าความน่าจะเป็นแบบทวินามได้ด้วยความน่าจะเป็นแบบปกติ โดยที่ $\mu = np = (200)(0.43) = 86$ และ $\sigma^2 = np(1-p) = (200)(0.43)(1-0.43) = 49.02$ นั่นคือ $\sigma = \sqrt{49.02}$ แต่ก่อนที่จะใช้การประมาณค่าด้วยความน่าจะเป็นแบบปกติ จะทำการปรับค่าด้วย 0.5 ก่อน

(1) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ตั้งแต่ 40 ถึง 100 คน แทนได้ด้วยสัญลักษณ์ $P(40 \le X \le 100)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(40 \le X \le 100) &= P(39.5 \le X \le 100.5) \\ &= P\Big(\frac{39.5-86}{\sqrt{49.02}} \le X \le \frac{100.5-86}{\sqrt{49.02}}\Big) \\ &= P(-6.64 \le Z \le 2.07) \\ &= 0.9808 \end{aligned}$$

ดังนั้น จะนักศึกษาจบที่ภายใน 4 ปี ตั้งแต่ 40 ถึง 100 คน ด้วยความน่าจะเป็นเท่ากับ 0.9808

(2) ความน่าจะเป็นที่จะมีนักศึกษาจบภายใน 4 ปี ระหว่าง 75 กับ 95 คน แทนได้ด้วยสัญลักษณ์ $P(75 < X < 95)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(75 < X < 95) &= P(75.5 \le X \le 94.5) \\ &= P\Big(\frac{75.5-86}{\sqrt{49.02}} \le X \le \frac{94.5-86}{\sqrt{49.02}}\Big) \\ &= P(-1.5 \le Z \le 1.21) \\ &= 0.8201 \end{aligned}$$

ดังนั้น จะมีนักศึกษาที่จบภายใน 4 ปี ระหว่าง 75 กับ 95 คน ด้วยความน่าจะเป็นเท่ากับ 0.8201 ■

ข้อสังเกต

จากตัวอย่างที่ 4.2.9 และ 4.2.10 จะเห็นว่าการประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปกติโดยมีการปรับค่าด้วย 0.5 (วิธีที่ 3) มีค่าใกล้เคียงกับการคำนวณโดยตรงจากฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทวินาม (วิธีที่ 1) มากกว่าการประมาณค่าความน่าจะเป็นแบบทวินามด้วยความน่าจะเป็นแบบปกติโดยไม่มีการปรับค่าด้วย 0.5 (วิธีที่ 2)

4.3 การแจกแจงความน่าจะเป็นแบบไคสแควร์

นิยามที่ 4.3.1 ถ้า $X_1, X_2, \ldots, X_\nu$ เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐานและเป็นอิสระต่อกันแล้ว

$$X = \sum_{i=1}^{\nu} X_i^2$$

จะเรียกว่าเป็นตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบไคสแควร์ (Chi-square probability distribution)*** ด้วยองศาอิสระ (degree of freedom) เท่ากับ ν

ถ้าตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสเท่ากับ ν แล้ว เราจะเขียนแทนได้ด้วยสัญลักษณ์

$$X \sim \chi^2(\nu)$$

หมายเหตุ

1. ถ้าตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 แล้ว $Z = \dfrac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$ ดังนั้น

$$Z^2 = \frac{(X-\mu)^2}{\sigma^2}$$

จะเรียกว่าเป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ 1 นั่นคือ

$$Z^2 \sim \chi^2(1)$$

2. ถ้าตัวแปรสุ่ม X_1 และ X_2 เป็นอิสระต่อกันและมีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ ν_1 และ ν_2 ตามลำดับ แล้วตัวแปรสุ่ม $X_1 + X_2$ จะมีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ $\nu_1 + \nu_2$ นั่นคือ

$$X_1 + X_2 \sim \chi^2(\nu_1 + \nu_2)$$

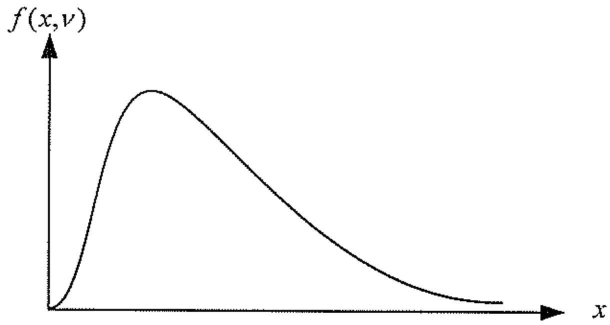
ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์

ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ ด้วยองศาอิสระเท่ากับ ν แล้ว ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการ

$$f(x,\nu) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} e^{-x/2} x^{(\frac{\nu}{2}-1)}, & x > 0 \\ 0, & x \le 0 \end{cases}$$

โดยที่ $\Gamma(\nu) = \int_0^\infty e^{-x} x^{\nu-1} dx$

สำหรับฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์แสดงได้ในรูปที่ 4.3.1



รูปที่ 4.3.1 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์

<u>หมายเหตุ</u>

ในหัวข้อนี้ใช้สัญลักษณ์ $f(x,\nu)$ แทน $f(x)$ เนื่องจากจะแสดงให้เห็นว่าฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์มีความสัมพันธ์กับองศาอิสระ

ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์

พิจารณา $X \sim \chi^2(\nu)$ จากนิยามค่าคาดหมายหรือค่าเฉลี่ยของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x f(x,\nu)dx \\
&= \int_{0}^{\infty} x \cdot \frac{1}{2^{\nu/2}\Gamma(\nu/2)} e^{-x/2} x^{(\frac{\nu}{2}-1)} dx \\
&= \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_{0}^{\infty} e^{-x/2} x^{\frac{\nu}{2}} dx
\end{aligned}
$$

กำหนดให้ $y = \frac{x}{2}$ จะได้ว่า $dy = \frac{1}{2}dx$ ดังนั้น

$$
\begin{aligned}
E(X) &= \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_{0}^{\infty} e^{-y} (2y)^{\frac{\nu}{2}} 2dy \\
&= \frac{2^{(\nu/2)+1}}{2^{\nu/2}\Gamma(\nu/2)} \int_{0}^{\infty} e^{-y} y^{\frac{\nu}{2}} dy \\
&= \frac{2}{\Gamma(\nu/2)} \int_{0}^{\infty} e^{-y} y^{\frac{\nu}{2}} dy
\end{aligned}
$$

พิจารณาการหาค่าด้วยการอินทิเกรตแบบแยกส่วน (by parts) โดยให้ $u = y^{\nu/2}$ และ $dv = e^{-y}$

$$
E(X) = \frac{2}{\Gamma(\nu/2)} \left(-y^{\nu/2} e^{-y} \Big|_{y=0}^{y=\infty} + \int_{0}^{\infty} e^{-y} \frac{\nu}{2} y^{\frac{\nu}{2}-1} dy \right)
$$

เนื่องจาก $\lim_{y\to\infty} y^{\nu/2}e^{-y} = \lim_{y\to 0} y^{\nu/2}e^{-y} = 0$ ดังนั้น

$$E(X) = \frac{2}{\Gamma(\nu/2)}\left(0 + \frac{\nu}{2}\int_0^\infty e^{-y}y^{\frac{\nu}{2}-1}dy\right)$$

เนื่องจาก $\Gamma(\nu/2) = \int_0^\infty e^{-x}x^{\frac{\nu}{2}-1}dx$ ดังนั้น

$$E(X) = \frac{2}{\Gamma(\nu/2)}\left(\frac{\nu}{2}\Gamma(\nu/2)\right) = \nu$$

ดังนั้น ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์มีค่าเท่ากับ

$$E(X) = \nu$$

ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์

กำหนดให้ $X \sim \chi^2(\nu)$ พิจารณา

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x,\nu)dx \\ &= \int_0^\infty x^2 \cdot \frac{1}{2^{\nu/2}\Gamma(\nu/2)} e^{-x/2}x^{(\frac{\nu}{2}-1)}dx \\ &= \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\int_0^\infty e^{-x/2}x^{\frac{\nu}{2}+1}dx \end{aligned}$$

กำหนดให้ $y = \frac{x}{2}$ จะได้ว่า $dy = \frac{1}{2}dx$ ดังนั้น

$$\begin{aligned} E(X^2) &= \frac{1}{2^{\nu/2}\Gamma(\nu/2)}\int_0^\infty e^{-y}(2y)^{\frac{\nu}{2}+1}2dy \\ &= \frac{2^{(\nu/2)+2}}{2^{\nu/2}\Gamma(\nu/2)}\int_0^\infty e^{-y}y^{\frac{\nu}{2}+1}dy \\ &= \frac{2^2}{\Gamma(\nu/2)}\int_0^\infty e^{-y}y^{\frac{\nu}{2}+1}dy \end{aligned}$$

พิจารณาการหาค่าด้วยการอินทิเกรตแบบแยกส่วน (by parts) โดยให้ $u = y^{(\nu/2)+1}$ และ $dv = e^{-y}$

$$E(X^2) = \frac{2^2}{\Gamma(\nu/2)}\left(-y^{(\nu/2)+1}e^{-y}\Big|_{y=0}^{y=\infty} + \int_0^\infty e^{-y}\left(\frac{\nu}{2}+1\right)y^{\frac{\nu}{2}}dy\right)$$

เนื่องจาก $\lim_{y\to\infty} y^{(\nu/2)+1}e^{-y} = \lim_{y\to 0} y^{(\nu/2)+1}e^{-y} = 0$ ดังนั้น

$$E(X^2) = \frac{2^2}{\Gamma(\nu/2)}\left(0 + \left(\frac{\nu}{2}+1\right)\int_0^\infty e^{-y}y^{\frac{\nu}{2}}dy\right)$$

พิจารณาการหาค่าด้วยการอินทิเกรตแบบแยกส่วน (by parts) อีกครั้ง โดยให้ $u = y^{\nu/2}$ และ $dv = e^{-y}$

$$E(X^2)=\frac{2^2}{\Gamma(\nu/2)}\left(\frac{\nu}{2}+1\right)\left(-y^{\nu/2}e^{-y}\Big|_{y=0}^{y=\infty}+\int_0^\infty e^{-y}\frac{\nu}{2}y^{\frac{\nu}{2}-1}dy\right)$$

เนื่องจาก $\lim_{y\to\infty} y^{\nu/2}e^{-y}=\lim_{y\to 0} y^{\nu/2}e^{-y}=0$ ดังนั้น

$$E(X^2)=\frac{2^2}{\Gamma(\nu/2)}\left(\frac{\nu}{2}+1\right)\left(0+\frac{\nu}{2}\int_0^\infty e^{-y}y^{\frac{\nu}{2}-1}dy\right)$$

เนื่องจาก $\Gamma(\nu/2)=\int_0^\infty e^{-x}x^{\frac{\nu}{2}-1}dx$ ดังนั้น

$$E(X^2)=\frac{2^2}{\Gamma(\nu/2)}\left(\frac{\nu}{2}+1\right)\left(\frac{\nu}{2}\right)\Gamma(\nu/2)=\nu^2+2\nu$$

จากนิยามความแปรปรวนของตัวแปรสุ่มจะได้ว่า

$$V(X)=E(X^2)-[E(X)]^2$$

นั่นคือ

$$V(X)=(\nu^2+2\nu)-\nu^2=2\nu$$

ดังนั้น ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์มีค่าเท่ากับ

$$V(X)=2\nu$$

สมบัติที่สำคัญของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์

1. กราฟมีลักษณะเบ้ขวาตามรูปที่ 4.3.1 และการเบ้จะขึ้นอยู่กับองศาอิสระ
2. พื้นที่ใต้กราฟแต่เหนือแกน X มีค่าเท่ากับ 1

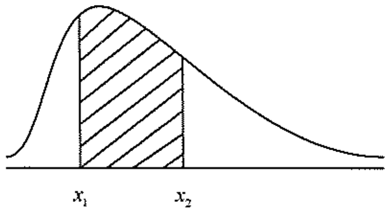
การหาค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์

พิจารณา $X\sim\chi^2(\nu)$ ถ้าต้องการหาค่าความน่าจะเป็นของตัวแปรสุ่ม X โดยที่ $x_1<X<x_2$ สามารถหาได้โดยใช้ฟังก์ชันความน่าจะเป็นดังนี้

$$P(x_1<X<x_2)=\int_{x_1}^{x_2}f(x,\nu)dx$$

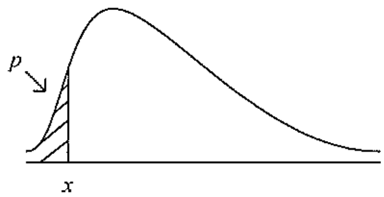
$$=\int_{x_1}^{x_2}\frac{1}{2^{\nu/2}\Gamma(\nu/2)}e^{-x/2}x^{(\frac{\nu}{2}-1)}dx$$

เนื่องจากฟังก์ชันความน่าจะเป็นมีค่ามากกว่าศูนย์เสมอ ดังนั้น $P(x_1<X<x_2)$ จึงมีค่าเท่ากับพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ โดยที่ $x_1<X<x_2$ นั่นเอง ตามรูปที่ 4.3.2



รูปที่ 4.3.2 พื้นที่ที่แรเงามีค่าเท่ากับความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ เมื่อ $x_1 < X < x_2$

การหาค่าความน่าจะเป็นโดยตรงจากการอินทิเกรตฟังก์ชันความน่าจะเป็นถือว่าค่อนข้างยุ่งยาก เพื่อความสะดวกในการหาค่าความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ จึงได้มีการจัดทำตารางเพื่อช่วยในการหาค่าความน่าจะเป็นสำหรับตัวแปรสุ่ม $X \sim \chi^2(\nu)$ พิจารณาตารางที่ 4 ในภาคผนวก ซึ่งจะให้ค่า x ที่ทำให้ $F(x) = P(X \le x) = p$ ดังรูปที่ 4.3.3



รูปที่ 4.3.3 พื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ เมื่อ $X \le x$ มีค่าเท่ากับ p

พิจารณาส่วนหนึ่งของตารางที่ 4.1 ดังนี้

ν	p									
	0.0005	0.001	0.005	0.01	0.025	0.05	0.10	0.20	$\boxed{0.30}$	0.40
$\boxed{6}$	0.299	0.381	0.676	0.872	1.237	1.635	2.204	3.070	$\boxed{3.828}$	4.570

มีความหมายว่า สำหรับตัวแปรสุ่ม X ใด ๆ ที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ 6 นั่นคือ $X \sim \chi^2(6)$

$$F(3.828) = P(X \le 3.828) = 0.30$$

การหาค่าความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ สามารถใช้สมบัติของกราฟฟังก์ชันความน่าจะเป็นมาช่วยได้ ตามตัวอย่างต่อไปนี้

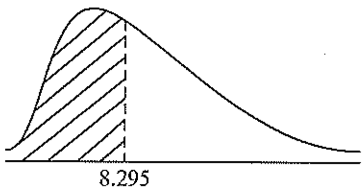
<u>ตัวอย่างที่ 4.3.1</u> กำหนดให้ X มีการแจกแจงแบบไคสแควร์ด้วยองศาอิสระเท่ากับ 10 นั่นคือ $X \sim \chi^2(10)$ จงหา

(1) $P(X < 8.295)$

(2) $P(X > 4.865)$

(3) $P(2.558 \le X \le 18.307)$

วิธีทำ (1) พิจารณา $P(X < 8.295)$ จะมีค่าเท่ากับพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ โดยที่ $X < 8.295$ ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป

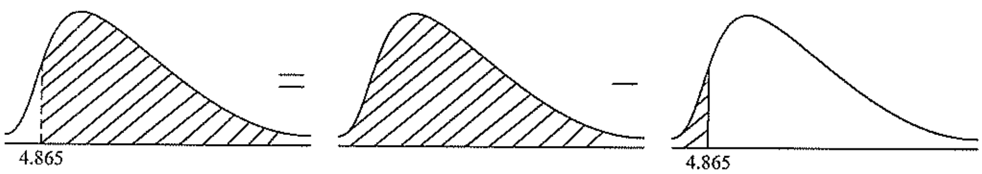


โดยใช้ตารางที่ 4.1 เมื่อ $v = 10$

$$P(X < 8.295) = 0.4$$

ดังนั้น $X < 8.295$ ด้วยความน่าจะเป็นเท่ากับ 0.4

(2) พิจารณา $P(X > 4.865)$ จะมีค่าเท่ากับพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ โดยที่ $X > 4.865$ เนื่องจากพื้นที่ใต้กราฟมีค่าเท่ากับ 1 ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป



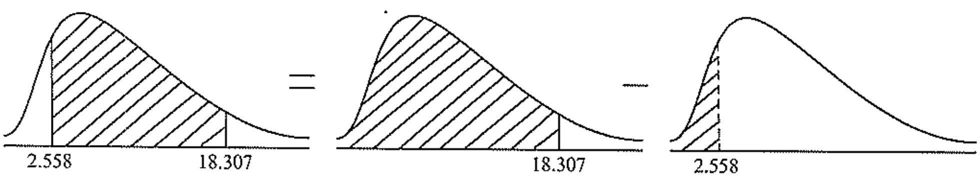
โดยใช้ตารางที่ 4.1 เมื่อ $v = 10$

$$P(X > 4.865) = 1 - P(X \le 4.865)$$

$$= 1 - 0.1 = 0.9$$

ดังนั้น $X > 4.865$ ด้วยความน่าจะเป็นเท่ากับ 0.9

(3) พิจารณา $P(2.558 \le X \le 18.307)$ จะมีค่าเท่ากับพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ โดยที่ $2.558 \le X \le 18.307$ ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป



โดยใช้ตารางที่ 4.1 และ 4.2 เมื่อ $\nu = 10$

$$P(2.558 \le X \le 18.307) = P(X \le 18.307) - P(X < 2.558)$$
$$= 0.95 - 0.01 = 0.94$$

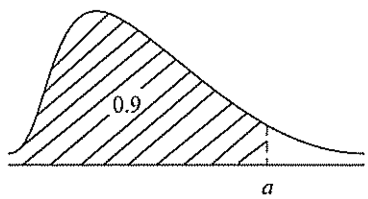
ดังนั้น $2.558 \le X \le 18.307$ ด้วยความน่าจะเป็นเท่ากับ 0.94 ■

<u>ตัวอย่างที่ 4.3.2</u> กำหนดให้ X มีการแจกแจงแบบไคสแควร์ด้วยองศาอิสระเท่ากับ 10 นั่นคือ $X \sim \chi^2(10)$ จงหา

(1) ค่า a ที่ทำให้ $P(X < a) = 0.9$

(2) ค่า a ที่ทำให้ $P(X > a) = 0.95$

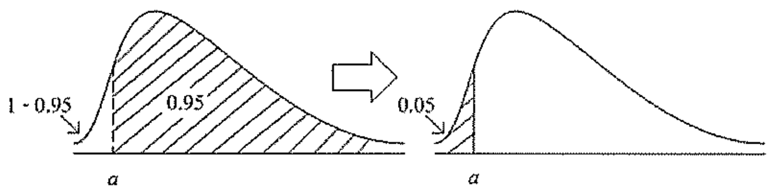
วิธีทำ (1) พิจารณา $P(X < a) = 0.9$ หมายความว่า พื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ โดยที่ $X < a$ มีค่าเท่ากับ 0.9 ดังรูป



จากตารางที่ 4.2 เมื่อ $\nu = 10$ จะได้ว่า $P(X < 15.987) = 0.9$

ดังนั้น ค่า a ที่ทำให้ $P(X < a) = 0.9$ มีค่าเท่ากับ 15.987

(2) พิจารณา $P(X > a) = 0.95$ หมายความว่า พื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ โดยที่ $X > a$ มีค่าเท่ากับ 0.95 เนื่องจากพื้นที่ใต้กราฟมีค่าเท่ากับ 1 จึงพิจารณาหาค่า a ได้ดังรูป



จากรูป $P(X \le a) = 0.05$ และจากตารางที่ 4.1 เมื่อ $\nu = 10$ จะได้ว่า $P(X \le 3.94) = 0.05$

ดังนั้น ค่า a ที่ทำให้ $P(X > a) = 0.95$ มีค่าเท่ากับ 3.94 ■

4.4 การแจกแจงความน่าจะเป็นแบบที

<u>นิยามที่ 4.4.1</u> ถ้า $Z \sim \mathcal{N}(0,1)$ และ $V \sim \chi^2(\nu)$ เป็นตัวแปรสุ่มที่เป็นอิสระต่อกันแล้ว

$$T = \frac{Z}{\sqrt{V/\nu}}$$

จะเรียกว่าเป็นตัวแปรสุ่มที่มี ***การแจกแจงความน่าจะเป็นแบบที (t-probability distribution or Student's t probability distribution)*** ด้วยองศาอิสระ (degree of freedom) เท่ากับ ν

ถ้าตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ ν แล้ว เราจะเขียนแทนได้ด้วยสัญลักษณ์

$$X \sim t(\nu)$$

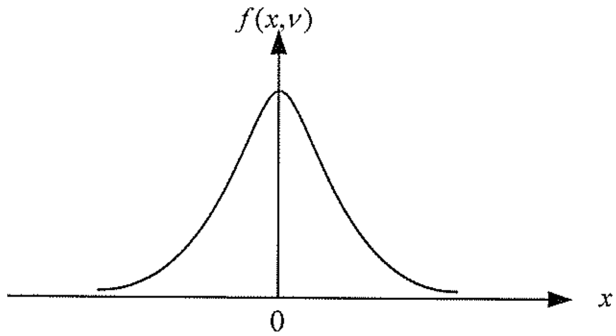
ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที

ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ ν แล้ว ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม X เป็นไปตามสมการ

$$f(x,\nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}$$

เมื่อ $-\infty < x < \infty$ โดยที่ $\Gamma(\nu) = \int_0^\infty e^{-x}x^{\nu-1}dx$

สำหรับฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีแสดงได้ในรูปที่ 4.4.1



รูปที่ 4.4.1 ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที

<u>หมายเหตุ</u>

ในหัวข้อนี้ใช้สัญลักษณ์ $f(x,\nu)$ แทน $f(x)$ เนื่องจากจะแสดงให้เห็นว่าฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีมีความสัมพันธ์กับองศาอิสระ

ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที

พิจารณา $X \sim t(\nu)$ จากนิยามค่าคาดหมายหรือค่าเฉลี่ยของตัวแปรสุ่มแบบต่อเนื่องจะได้ว่า

$$E(X)=\int_{-\infty}^{\infty} xf(x,\nu)dx$$

พิจารณา

$$\int_0^{\infty} xf(x,\nu)dx=\int_0^{\infty} x\cdot\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}dx$$

$$=\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\int_0^{\infty} x\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}dx$$

กำหนดให้ $y=1+\frac{x^2}{\nu}$ จะได้ว่า $dy=\frac{2}{\nu}xdx$ ดังนั้น

$$\int_0^{\infty} xf(x,\nu)dx=\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\int_1^{\infty} y^{-(\nu+1)/2}\frac{\nu}{2}dy$$

$$=\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\cdot\frac{\nu}{2}\int_1^{\infty} y^{-(\nu+1)/2}dy$$

$$=\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\cdot\frac{\nu}{2}\left[\frac{y^{1-(\nu+1)/2}}{1-(\nu+1)/2}\right]_{y=1}^{y=\infty}$$

$$=\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\cdot\frac{\nu}{2}\cdot\frac{2}{1-\nu}\left[\frac{1}{y^{-(1-\nu)/2}}\right]_{y=1}^{y=\infty}$$

ถ้า $\nu>1$ แล้ว $\lim\limits_{y\to\infty}\frac{1}{y^{-(1-\nu)/2}}=0$ ดังนั้น

$$\int_0^{\infty} xf(x,\nu)dx=-\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\cdot\frac{\nu}{2}\cdot\frac{2}{1-\nu}$$

$$=\frac{\nu}{\nu-1}\cdot\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}<\infty$$

และเนื่องจาก $\int_{-\infty}^{\infty} xf(x,\nu)dx=\frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\int_0^{\infty} x\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2}dx$ เป็นฟังก์ชันคี่ ดังนั้น

$$\int_{-\infty}^{\infty} xf(x,\nu)dx=0$$

ดังนั้น เมื่อ $\nu>1$

$$E(X)=\int_{-\infty}^{\infty} xf(x,\nu)dx=0$$

สำหรับ $\nu>1$ ดังนั้น ค่าเฉลี่ยของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีมีค่าเท่ากับ

$$E(X)=0$$

ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที

กำหนดให้ $X \sim t(\nu)$ พิจารณา

$$\int_0^\infty x^2 f(x,\nu)dx = \int_0^\infty x^2 \cdot \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} dx$$

$$= \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\int_0^\infty x\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} dx$$

กำหนดให้ $y = \frac{x^2}{\nu}$ จะได้ว่า $dy = \frac{2}{\nu}xdx$ ดังนั้น

$$\int_0^\infty x^2 f(x,\nu)dx = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\int_0^\infty y^{1/2}(1+y)^{-(\nu+1)/2}\frac{\nu^{3/2}}{2}dy$$

$$= \frac{\nu^{3/2}\Gamma((\nu+1)/2)}{2\Gamma(\nu/2)\sqrt{\pi\nu}}\int_0^\infty y^{1/2}(1+y)^{-(\nu+1)/2}dy$$

$$= \frac{\nu^{3/2}\Gamma((\nu+1)/2)}{2\Gamma(\nu/2)\sqrt{\pi\nu}}\int_0^\infty y^{\frac{3}{2}-1}(1+y)^{-\frac{3}{2}-(\frac{\nu}{2}-1)}dy$$

เนื่องจาก $B(x,y) = \int_0^\infty t^{x-1}(1+t)^{-x-y}dt$ จะลู่เข้าเมื่อ $x > 0$ และ $y > 0$ ดังนั้น $\frac{\nu}{2} - 1 > 0$ หรือ $\nu > 2$ นั่นคือ สำหรับกรณีที่ $\nu > 2$

$$\int_0^\infty x^2 f(x,\nu)dx = \frac{\nu^{3/2}\Gamma((\nu+1)/2)}{2\Gamma(\nu/2)\sqrt{\pi\nu}}B(\frac{3}{2},\frac{\nu}{2}-1)$$

และจากความสัมพันธ์ระหว่างฟังก์ชันแกมมาและฟังก์ชันเบต้าดังนี้ $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ จะได้ว่า

$$\int_0^\infty x^2 f(x,\nu)dx = \frac{\nu^{3/2}\Gamma((\nu+1)/2)}{2\Gamma(\nu/2)\sqrt{\pi\nu}}\cdot\frac{\Gamma(\frac{3}{2})\Gamma(\frac{\nu}{2}-1)}{\Gamma(\frac{\nu}{2}+\frac{1}{2})}$$

เนื่องจาก $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ จะได้ว่า

$$\int_0^\infty x^2 f(x,\nu)dx = \frac{\nu\Gamma((\nu+1)/2)}{2\Gamma(\nu/2)\Gamma(\frac{1}{2})}\cdot\frac{\Gamma(\frac{3}{2})\Gamma(\frac{\nu}{2}-1)}{\Gamma(\frac{\nu}{2}+\frac{1}{2})}$$

และจาก $\Gamma(\nu) = (\nu-1)\Gamma(\nu-1)$ จะได้ว่า

$$\int_0^\infty x^2 f(x,\nu)dx = \frac{\nu\Gamma(\frac{\nu}{2}+\frac{1}{2})}{2(\frac{\nu}{2}-1)\Gamma(\frac{\nu}{2}-1)\Gamma(\frac{1}{2})}\cdot\frac{\frac{1}{2}\Gamma(\frac{1}{2})\Gamma(\frac{\nu}{2}-1)}{\Gamma(\frac{\nu}{2}+\frac{1}{2})}$$

$$= \frac{\nu}{2(\nu-2)}$$

เนื่องจาก $x^2 f(x,\nu)$ เป็นฟังก์ชันคู่ ทำให้ได้ว่า

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x,\nu)dx$$

$$= 2\int_{0}^{\infty} x^2 f(x,\nu)dx$$

$$= \frac{\nu}{\nu-2}$$

จากนิยามความแปรปรวนของตัวแปรสุ่มจะได้ว่า

$$V(X) = E(X^2) - [E(X)]^2$$

แต่เนื่องจาก $E(X) = 0$ จึงได้ว่า

$$V(X) = E(X^2)$$

$$= \frac{\nu}{\nu-2}$$

ดังนั้น สำหรับ $\nu > 2$ ความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีมีค่าเท่ากับ

$$V(X) = \frac{\nu}{\nu-2}$$

สมบัติที่สำคัญของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที

1. กราฟเป็นรูประฆังคว่ำตามรูปที่ 4.4.1
2. กราฟมีลักษณะสมมาตรตามแนวแกน Y ที่ $X = 0$ ตามรูปที่ 4.4.1
3. พื้นที่ใต้กราฟแต่เหนือแกน X มีค่าเท่ากับ 1

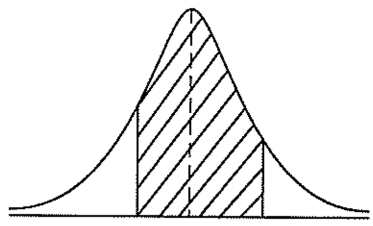
การหาค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที

พิจารณา $X \sim t(\nu)$ ถ้าต้องการหาค่าความน่าจะเป็นของตัวแปรสุ่ม X โดยที่ $x_1 < X < x_2$ สามารถหาได้โดยใช้ฟังก์ชันความน่าจะเป็นดังนี้

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x,\nu)dx$$

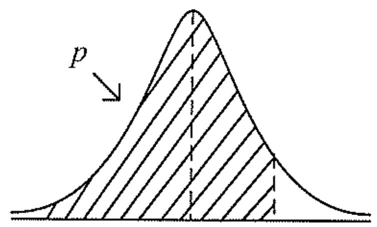
$$= \int_{x_1}^{x_2} \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\frac{x^2}{\nu}\right)^{-(\nu+1)/2} dx$$

เนื่องจากฟังก์ชันความน่าจะเป็นมีค่ามากกว่าศูนย์เสมอ ดังนั้น $P(x_1 < X < x_2)$ จึงมีค่าเท่ากับพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที โดยที่ $x_1 < X < x_2$ นั่นเอง ตามรูปที่ 4.4.2



รูปที่ 4.4.2 พื้นที่ที่แรเงามีค่าเท่ากับความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที เมื่อ $x_1 < X < x_2$

การหาค่าความน่าจะเป็นโดยตรงจากการอินทิเกรตฟังก์ชันความน่าจะเป็นถือว่าค่อนข้างยุ่งยาก เพื่อความสะดวกในการหาค่าความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที จึงได้มีการจัดทำตารางเพื่อช่วยในการหาค่าความน่าจะเป็นสำหรับตัวแปรสุ่ม $X \sim t(\nu)$ พิจารณาตารางที่ 5 ในภาคผนวก ซึ่งจะให้ค่า x ที่ทำให้ $F(x) = P(X \leq x) = p$ ดังรูปที่ 4.4.3



รูปที่ 4.4.3 พื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที เมื่อ $X \leq x$ มีค่าเท่ากับ p

พิจารณาส่วนหนึ่งของตารางที่ 5 ดังนี้

ν	p									
	0.60	0.70	0.80	0.90	$\boxed{0.95}$	0.975	0.990	0.995	0.999	0.9995
$\boxed{3}$	~~0.277~~	~~0.584~~	~~0.978~~	~~1.638~~	$\boxed{2.353}$	3.182	4.541	5.841	10.215	12.924

มีความหมายว่า สำหรับตัวแปรสุ่ม X ใด ๆ ที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ 3 นั่นคือ $X \sim t(3)$

$$F(2.353) = P(X \leq 2.353) = 0.95$$

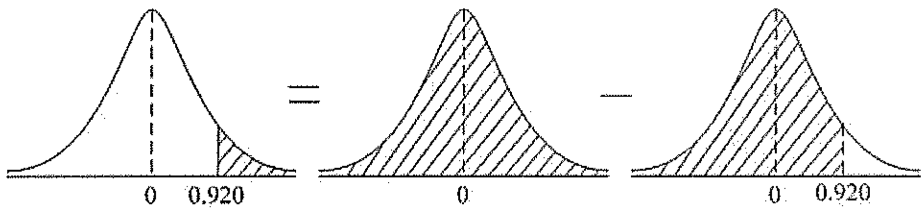
การหาค่าความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที สามารถใช้สมบัติของกราฟฟังก์ชันความน่าจะเป็นมาช่วยได้ ตามตัวอย่างต่อไปนี้

<u>ตัวอย่างที่ 4.4.1</u> กำหนดให้ X มีการแจกแจงแบบทีด้วยองศาอิสระเท่ากับ 5 นั่นคือ $X \sim t(5)$ จงหา

(1) $P(X \geq 0.920)$

(2) $P(-1.476 \le X \le 2.571)$

วิธีทำ (1) พิจารณา $P(X \ge 0.920)$ จะมีค่าเท่ากับพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที โดยที่ $X \ge 0.920$ เนื่องจากพื้นที่ใต้กราฟมีค่าเท่ากับ 1 ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป

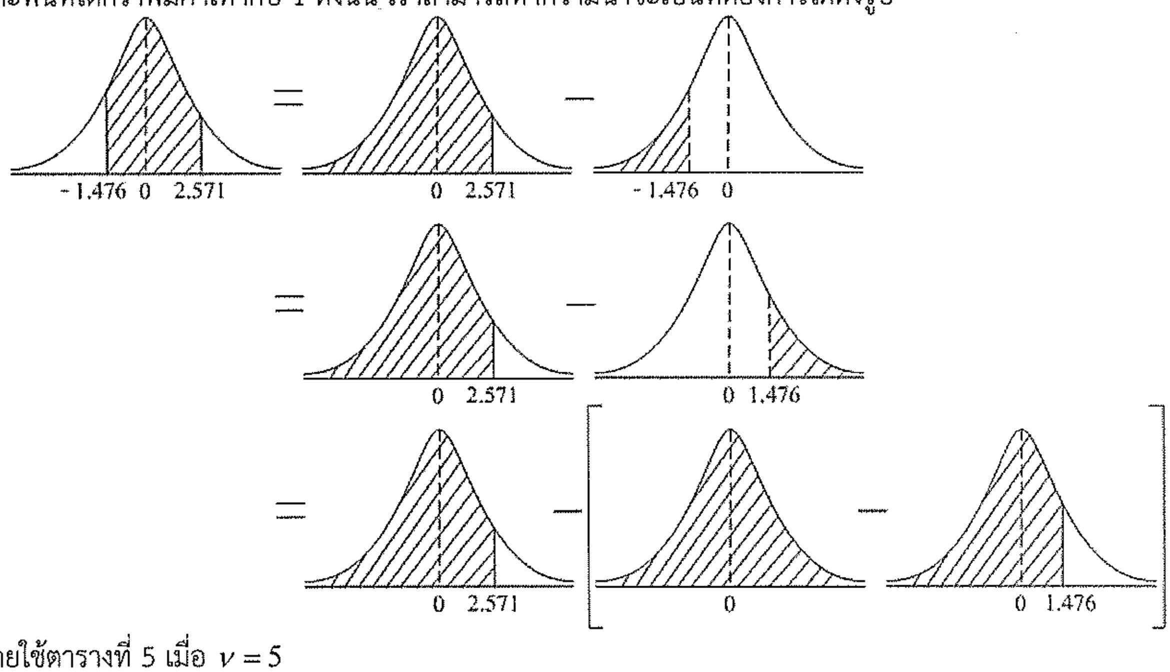


โดยใช้ตารางที่ 5 เมื่อ $v = 5$

$$P(X \ge 0.920) = 1 - P(X < 0.920)$$
$$= 1 - 0.80 = 0.20$$

ดังนั้น $X \ge 0.920$ ด้วยความน่าจะเป็นเท่ากับ 0.20

(2) พิจารณา $P(-1.476 \le X \le 2.571)$ จะมีค่าเท่ากับพื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที โดยที่ $-1.476 \le X \le 2.571$ เนื่องจากกราฟมีสมบัติสมมาตร และพื้นที่ใต้กราฟมีค่าเท่ากับ 1 ดังนั้น เราสามารถหาความน่าจะเป็นที่ต้องการได้ดังรูป



โดยใช้ตารางที่ 5 เมื่อ $v = 5$

$$P(-1.476 \le X \le 2.571) = P(X \le 2.571) - P(X < -1.476)$$
$$= P(X \le 2.571) - P(X > 1.476)$$
$$= P(X \le 2.571) - \{1 - P(X \le 1.476)\}$$

$$= 0.975 - \{1 - 0.90\} = 0.875$$

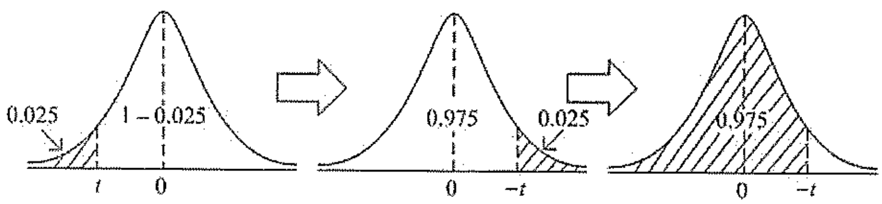
ดังนั้น $-1.476 \le X \le 2.571$ ด้วยความน่าจะเป็นเท่ากับ 0.875 ■

<u>ตัวอย่างที่ 4.4.2</u> กำหนดให้ X มีการแจกแจงแบบทีด้วยองศาอิสระเท่ากับ 5 นั่นคือ $X \sim t(5)$ จงหา

(1) ค่า t ที่ทำให้ $P(X < t) = 0.025$

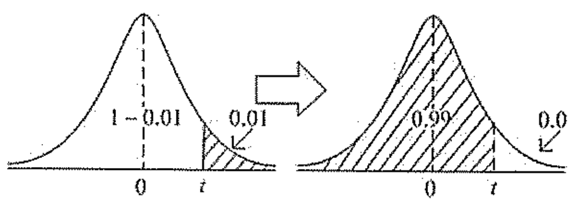
(2) ค่า t ที่ทำให้ $P(X > t) = 0.01$

วิธีทำ (1) พิจารณา $P(X < t) = 0.025$ หมายความว่า พื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที โดยที่ $X < t$ มีค่าเท่ากับ 0.025 แต่เนื่องจากพื้นที่ที่กำหนดให้มีค่าน้อยกว่า 0.5 ดังนั้น $t < 0$ โดยสมบัติสมมาตร และพื้นที่ใต้กราฟมีค่าเท่ากับ 1 จึงพิจารณาการหาค่า t ได้ดังรูป



จากรูป $P(X \le -t) = 0.975$ และจากตารางที่ 5 เมื่อ $\nu = 5$ จะได้ว่า $P(X \le 2.571) = 0.975$
ดังนั้น ค่า t ที่ทำให้ $P(X < t) = 0.025$ มีค่าเท่ากับ − 2.571

(2) พิจารณา $P(X > t) = 0.01$ หมายความว่า พื้นที่ใต้กราฟของฟังก์ชันความน่าจะเป็นสำหรับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที โดยที่ $X > t$ มีค่าเท่ากับ 0.01 แต่เนื่องจากพื้นที่ที่กำหนดให้มีค่าน้อยกว่า 0.5 ดังนั้น $t > 0$ และจากพื้นที่ใต้กราฟมีค่าเท่ากับ 1 จึงพิจารณาการหาค่า t ได้ดังรูป



จากรูป $P(X \le t) = 0.99$ และจากตารางที่ 5 เมื่อ $\nu = 5$ จะได้ว่า $P(X \le 3.365) = 0.99$
ดังนั้น ค่า t ที่ทำให้ $P(X > t) = 0.01$ มีค่าเท่ากับ 3.365 ■

บทที่ 5
การแจกแจงของฟังก์ชันที่ได้จากกลุ่มตัวอย่าง
(Sampling Distribution)

ในบทนี้จะกล่าวถึงการแจกแจงของฟังก์ชันที่ได้จากกลุ่มตัวอย่างซึ่งจะนำไปสู่การอนุมานทางสถิติ (statistical inference) หรือการนำข้อมูลของกลุ่มตัวอย่างไปสรุปผลเกี่ยวกับประชากร โดยที่การอนุมานทางสถิติจะกล่าวถึงอีกครั้งในบทที่ 6 และ 7 ต่อไป

5.1 การสุ่มตัวอย่างและฟังก์ชันต่าง ๆ ที่ได้จากกลุ่มตัวอย่าง

เนื่องจากการพิจารณาข้อมูลทั้งหมดที่เราต้องการทราบ บางครั้งในทางปฏิบัติอาจจะใช้เวลามาก หรืออาจเสียค่าใช้จ่ายสูงในการรวบรวมข้อมูล หรืออาจทำให้ข้อมูลไม่ทันสมัย ดังนั้นจึงต้องมีการเลือกเพียงบางส่วนของข้อมูลเพื่อนำมาวิเคราะห์แล้วสรุปผลไปยังสิ่งที่ต้องการทราบแทน ซึ่งเป็นที่มาที่สำคัญของการสุ่มตัวอย่าง เริ่มต้นหัวข้อนี้ด้วยนิยามที่สำคัญเกี่ยวกับการสุ่มตัวอย่าง

นิยามที่ 5.1.1

การสุ่มตัวอย่าง (random sampling) หมายถึง การเลือกข้อมูลบางส่วนเพื่อมาเป็นตัวแทนของข้อมูลทั้งหมดที่สนใจ

ประชากร (population) หมายถึง ข้อมูลทั้งหมดเกี่ยวกับสิ่งที่สนใจ

กลุ่มตัวอย่าง (sample) หมายถึง ส่วนหนึ่งของข้อมูลที่ถูกเลือกมาเป็นตัวแทนของข้อมูล

พารามิเตอร์ (parameter) หมายถึง ค่าที่บอกถึงลักษณะที่สำคัญของประชากร หรือค่าที่ประมวลได้จากประชากร

ค่าสถิติ (statistic) หมายถึง ค่าที่บอกถึงลักษณะที่สำคัญของกลุ่มตัวอย่าง หรือค่าที่ประมวลได้จากกลุ่มตัวอย่าง

ตัวอย่างที่ 5.1.1 เมื่อต้องการศึกษาข้อมูลเกี่ยวกับน้ำหนักของเด็กไทย ดังนั้น ประชากร คือ น้ำหนักของเด็กไทยทั้งประเทศ ส่วนกลุ่มตัวอย่าง คือ น้ำหนักของเด็กไทยที่สุ่มมาทำการศึกษา ถ้าทำการหาค่าเฉลี่ยและความแปรปรวนของน้ำหนักเด็กไทย ดังนั้น พารามิเตอร์ คือ ค่าเฉลี่ยและความแปรปรวนของน้ำหนักเด็กไทยทั้งประเทศ ส่วนค่าสถิติ คือ ค่าเฉลี่ยและความแปรปรวนของน้ำหนักเด็กไทยที่สุ่มมาทำการศึกษา

วิธีทำ ถ้าต้องการเลือกบุตรจากครอบครัวนี้มา 4 คน เพื่อเป็นตัวแทนของบุตรทั้งหมด จะสามารถเลือกได้ทั้งหมด 15 วิธี ดังนี้

วิธีที่	ชื่อ	น้ำหนัก	น้ำหนักเฉลี่ย	ความน่าจะเป็น
1	สมชาย สมหญิง สมทรง สมศักดิ์	50, 42, 42, 56	47.5	1/15
2	สมชาย สมหญิง สมทรง สมใจ	50, 42, 42, 40	43.5	1/15
3	สมชาย สมหญิง สมทรง สมคิด	50, 42, 42, 44	44.5	1/15
4	สมชาย สมหญิง สมศักดิ์ สมใจ	50, 42, 56, 40	47	1/15
5	สมชาย สมหญิง สมศักดิ์ สมคิด	50, 42, 56, 44	48	1/15
6	สมชาย สมหญิง สมใจ สมคิด	50, 42, 40, 44	44	1/15
7	สมชาย สมทรง สมศักดิ์ สมใจ	50, 42, 56, 40	47	1/15
8	สมชาย สมทรง สมศักดิ์ สมคิด	50, 42, 56, 44	48	1/15
9	สมชาย สมทรง สมใจ สมคิด	50, 42, 40, 44	44	1/15
10	สมชาย สมศักดิ์ สมใจ สมคิด	50, 56, 40, 44	47.5	1/15
11	สมหญิง สมทรง สมศักดิ์ สมใจ	42, 42, 56, 40	45	1/15
12	สมหญิง สมทรง สมศักดิ์ สมคิด	42, 42, 56, 44	46	1/15
13	สมหญิง สมทรง สมใจ สมคิด	42, 42, 40, 44	42	1/15
14	สมหญิง สมศักดิ์ สมใจ สมคิด	42, 56, 40, 44	45.5	1/15
15	สมทรง สมศักดิ์ สมใจ สมคิด	42, 56, 40, 44	45.5	1/15

จากตารางข้างต้นจะพบว่าค่าเฉลี่ยของกลุ่มตัวอย่างมีค่าที่เป็นไปได้ทั้งหมด 10 ค่า ซึ่งประกอบไปด้วย 42, 43.5, 44, 44.5, 45, 45.5, 46, 47, 47.5 และ 48 โดยที่แต่ละค่ามีความน่าจะเป็นตามตารางต่อไปนี้

น้ำหนักเฉลี่ย $\overline{x}_i$	ความถี่	ความน่าจะเป็น $P(\overline{X} = \overline{x}_i)$
42	1	1/15
43.5	1	1/15
44	2	2/15
44.5	1	1/15
45	1	1/15
45.5	2	2/15
46	1	1/15
47	2	2/15
47.5	2	2/15
48	2	2/15

(1) ค่าเฉลี่ยของตัวแปรสุ่ม X เขียนแทนได้ด้วยสัญลักษณ์ $E(X)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} E(X) &= \sum_{i=1}^{6} x_i f(x_i) \\ &= \sum_{i=1}^{6} x_i P(X = x_i) \\ &= 50\left(\frac{1}{6}\right)+\ldots+44\left(\frac{1}{6}\right) \\ &= \frac{274}{6} \\ &\approx 45.67 \end{aligned}$$

ดังนั้น ค่าเฉลี่ยของตัวแปรสุ่ม X มีค่าเท่ากับ 45.67

(2) ค่าเฉลี่ยของตัวแปรสุ่ม $\bar{X}$ เขียนแทนได้ด้วยสัญลักษณ์ $E(\bar{X})$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} E(\bar{X}) &= \sum_{i-1}^{10} \bar{x}_i f(\bar{x}_i) \\ &= \sum_{i-1}^{10} \bar{x}_i P(\bar{X} = \bar{x}_i) \\ &= 42\left(\frac{1}{15}\right)+\ldots+48\left(\frac{2}{15}\right) \\ &= \frac{685}{15} \\ &\approx 45.67 \end{aligned}$$

ดังนั้น ค่าเฉลี่ยของตัวแปรสุ่ม $\bar{X}$ มีค่าเท่ากับ 45.67

(3) ความแปรปรวนของตัวแปรสุ่ม X เขียนแทนได้ด้วยสัญลักษณ์ $V(X)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 \\ &= \sum_{i-1}^{n} x_i^2 f(x_i) - E(X)^2 \\ &= \sum_{i-1}^{n} x_i^2 P(X = x_i) - E(X)^2 \\ &= \left((50)^2\left(\frac{1}{6}\right)+\ldots+(44)^2\left(\frac{1}{6}\right)\right)-\left(\frac{274}{6}\right)^2 \\ &= \frac{12700}{6} - \frac{75076}{36} \\ &= \frac{1124}{36} \\ &\approx 31.22 \end{aligned}$$

ดังนั้น ความแปรปรวนของตัวแปรสุ่ม X มีค่าเท่ากับ 31.22

(4) ความแปรปรวนของตัวแปรสุ่ม $\bar{X}$ เขียนแทนได้ด้วยสัญลักษณ์ $V(\bar{X})$ ซึ่งหาค่าได้ดังนี้

$$
\begin{aligned}
V(\bar{X}) &= E(\bar{X}^2) - E(\bar{X})^2 \\
&= \sum_{i=1}^{n} \bar{x}_i^2 f(\bar{x}_i) - E(\bar{X})^2 \\
&= \sum_{i=1}^{n} \bar{x}_i^2 P(\bar{X} = \bar{x}_i) - E(\bar{X})^2 \\
&= \left((42)^2 \left(\frac{1}{15} \right) + \ldots + (48)^2 \left(\frac{2}{15} \right) \right) - \left(\frac{685}{15} \right)^2 \\
&= \frac{31328.5}{15} - \frac{469225}{225} \\
&= \frac{702.5}{225} \\
&\approx 3.12
\end{aligned}
$$

ดังนั้น ความแปรปรวนของตัวแปรสุ่ม $\bar{X}$ มีค่าเท่ากับ 3.12

<u>ข้อสังเกต</u>

1. จากตัวอย่างที่ 5.2.1 จะพบว่า $E(\bar{X}) = E(X)$ แต่ $V(\bar{X}) \neq V(X)$

2. จากตัวอย่างที่ 5.2.1 เป็นการสุ่มตัวอย่างแบบไม่ใส่คืน แต่ในทางปฏิบัติแล้วยังมีการสุ่มตัวอย่างอีกแบบหนึ่งคือการสุ่มแบบใส่คืน ซึ่งถ้าขนาดของกลุ่มตัวอย่างมีขนาดเล็กเมื่อเทียบกับขนาดของประชากรแล้ว จะถือว่าการสุ่มทั้งสองแบบแทบจะไม่มีความแตกต่างกัน

5.3 ทฤษฎีบทต่าง ๆ ที่สำคัญ

ในหัวข้อนี้ได้กล่าวถึงทฤษฎีบทต่าง ๆ ที่สำคัญที่นำไปใช้ในการแจกแจงของฟังก์ชันที่ได้จากกลุ่มตัวอย่าง ซึ่งบางทฤษฎีอาจไม่ได้แสดงการพิสูจน์เอาไว้เนื่องจากรายละเอียดอยู่นอกเหนือขอบเขตของหนังสือเล่มนี้

ทฤษฎีบทที่ 5.3.1 กำหนดให้ $X_1, \ldots, X_n$ เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน โดยที่ $X_i \sim N(\mu_i, \sigma_i^2)$ เมื่อ $i = 1, \ldots, n$ ดังนั้น ตัวแปรสุ่ม $Y = a_1 X_1 + \ldots + a_n X_n$ จะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ $a_1\mu_1 + \ldots + a_n\mu_n$ และความแปรปรวนเท่ากับ $a_1^2\sigma_1^2 + \ldots + a_n^2\sigma_n^2$ นั่นคือ

$$Y \sim N(a_1\mu_1 + \ldots + a_n\mu_n, a_1^2\sigma_1^2 + \ldots + a_n^2\sigma_n^2)$$

<u>หมายเหตุ</u>

1. ถ้า $X_i \sim N(\mu,\sigma^2)$ เมื่อ $i=1,\ldots,n$ ดังนั้น จากทฤษฎีบทที่ 5.3.1 จะได้ว่าตัวแปรสุ่ม $\bar{X}=\frac{X_1+\ldots+X_n}{n}$ จะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu,\frac{\sigma^2}{n})$$

2. ความแปรปรวนของประชากร (σ^2) มักไม่ทราบค่า ในทางปฏิบัติจึงนิยมใช้ความแปรปรวนของกลุ่มตัวอย่าง (s^2) แทน

ทฤษฎีบทที่ 5.3.2 ทฤษฎีบทลิมิตเข้าสู่ส่วนกลาง (central limit theorem) กำหนดให้ $X_1,\ldots,X_n$ เป็นกลุ่มตัวอย่างขนาด n ซึ่งถูกเลือกมาจากประชากรที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 ดังนั้น เมื่อ $n\to\infty$ จะได้ว่าค่าเฉลี่ยของกลุ่มตัวอย่างหรือ $\bar{X}=\frac{X_1+\ldots+X_n}{n}$ จะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu,\frac{\sigma^2}{n})$$

ทฤษฎีบทลิมิตเข้าสู่ส่วนกลางเป็นทฤษฎีบทหนึ่งที่สำคัญมากในทฤษฎีบทที่เกี่ยวข้องกับความน่าจะเป็น และทำให้สามารถอธิบายได้ว่าปรากฏการณ์ทางธรรมชาติส่วนใหญ่มีการแจกแจงความน่าจะเป็นใกล้เคียงกับการแจกแจงแบบปกติ

<u>หมายเหตุ</u>

1. จากทฤษฎีลิมิตเข้าสู่ส่วนกลาง เมื่อ $n\to\infty$ จะได้ว่าตัวแปรสุ่ม $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน นั่นคือ

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

2. จากทฤษฎีลิมิตเข้าสู่ส่วนกลางและทฤษฎีบทที่ 5.3.1 เมื่อ $n\to\infty$ จะได้ว่าตัวแปรสุ่ม $X_1+\ldots+X_n$ มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ $n\mu$ และความแปรปรวนเท่ากับ $n\sigma^2$ นั่นคือ

$$X_1+\ldots+X_n \sim N(n\mu,n\sigma^2)$$

3. จากทฤษฎีลิมิตเข้าสู่ส่วนกลางจะใช้ได้เมื่อ $n\to\infty$ แต่ในทางปฏิบัตินิยมใช้เมื่อ $n\geq 30$

4. ความแปรปรวนของประชากร (σ^2) มักไม่ทราบค่า ในทางปฏิบัติจึงนิยมใช้ความแปรปรวนของกลุ่มตัวอย่าง (s^2) แทน

ทฤษฎีบทที่ 5.3.3 กำหนดให้ $X_1,...,X_n$ เป็นกลุ่มตัวอย่างขนาด n ซึ่งถูกเลือกมาจากประชากรที่มีการแจกแจงความน่าจะเป็นแบบปกติที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 ถ้า s^2 เป็นความแปรปรวนของกลุ่มตัวอย่าง แล้วตัวแปรสุ่ม $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ จะมีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

<u>พิสูจน์</u> เนื่องจาก $X_1,...,X_n$ เป็นกลุ่มตัวอย่างขนาด n ซึ่งถูกเลือกมาจากประชากรที่มีการแจกแจงความน่าจะเป็นแบบปกติซึ่งมีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 ดังนั้น $X_i \sim N(\mu,\sigma^2)$ เมื่อ $i=1,\dots,n$ พิจารณา

$$\begin{aligned}\sum_{i=1}^{n}(X_i-\mu)^2 &= \sum_{i=1}^{n}[(X_i-\bar{X})+(\bar{X}-\mu)]^2 \\ &= \sum_{i=1}^{n}(X_i-\bar{X})^2 + 2(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\bar{X}) + \sum_{i=1}^{n}(\bar{X}-\mu)^2\end{aligned}$$

เนื่องจาก $\sum_{i=1}^{n}(X_i-\bar{X})=0$ ดังนั้น

$$\sum_{i=1}^{n}(X_i-\mu)^2 = \sum_{i=1}^{n}(X_i-\bar{X})^2 + n(\bar{X}-\mu)^2$$

จะได้ว่า

$$\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} = \sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{\sigma^2} + \frac{n(\bar{X}-\mu)^2}{\sigma^2}$$

และจาก $s^2 = \sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n-1}$ ดังนั้น

$$\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} = \frac{(n-1)s^2}{\sigma^2} + \frac{(\bar{X}-\mu)^2}{\sigma^2/n}$$

เนื่องจาก $\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}$ มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ n และ $\frac{(\bar{X}-\mu)^2}{\sigma^2/n}$ มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ 1 ดังนั้น $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

ทฤษฎีบทที่ 5.3.4 กำหนดให้ $X_1,\dots,X_n$ เป็นกลุ่มตัวอย่างขนาด n ซึ่งถูกเลือกมาจากประชากรที่มีการแจกแจงความน่าจะเป็นแบบปกติที่มีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 ถ้า $X_1,\dots,X_n$ เป็นอิสระต่อกัน แล้วตัวแปรสุ่ม $T=\dfrac{\bar{X}-\mu}{s/\sqrt{n}}$ จะมีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$T=\frac{\bar{X}-\mu}{s/\sqrt{n}}\sim t(n-1)$$

<u>พิสูจน์</u> เนื่องจาก $X_1,\dots,X_n$ เป็นกลุ่มตัวอย่างขนาด n ซึ่งถูกเลือกมาจากประชากรที่มีการแจกแจงความน่าจะเป็นแบบปกติซึ่งมีค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 ดังนั้น $X_i\sim N(\mu,\sigma^2)$ เมื่อ $i=1,\dots,n$ พิจารณา

$$\begin{aligned}
T&=\frac{\bar{X}-\mu}{s/\sqrt{n}}\\
&=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\cdot\sqrt{\frac{\sigma^2}{s^2}}\\
&=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\cdot\sqrt{\frac{\sigma^2(n-1)}{s^2(n-1)}}\\
&=Z\cdot\sqrt{\frac{n-1}{V}}\\
&=\frac{Z}{\sqrt{V/(n-1)}}
\end{aligned}$$

เมื่อ $Z=\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ และ $V=\dfrac{(n-1)s^2}{\sigma^2}$

เนื่องจาก Z มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ 1 ส่วน V มีการแจกแจงแบบไคสแควร์ด้วยองศาอิสระเท่ากับ $n-1$ โดยที่ Z และ V เป็นอิสระต่อกัน ดังนั้น $T=\dfrac{\bar{X}-\mu}{s/\sqrt{n}}$ จึงมีการแจกแจงแบบทีด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$T=\frac{\bar{X}-\mu}{s/\sqrt{n}}\sim t(n-1)$$

ทฤษฎีบทที่ 5.3.5 กำหนดให้ตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบทวินาม ดังนี้ $X \sim B(n,p)$ ถ้า $n \to \infty$ ค่าสัดส่วนของกลุ่มตัวอย่าง $\hat{p} = \frac{x}{n}$ จะมีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเท่ากับ p และความแปรปรวนเท่ากับ $\frac{p(1-p)}{n}$ นั่นคือ เมื่อ $n \to \infty$

$$\hat{p} \sim N(p, \frac{p(1-p)}{n})$$

<u>พิสูจน์</u> พิจารณาตัวแปรสุ่ม

$$\hat{p} = \frac{x}{n} = \frac{X_1 + \ldots + X_n}{n}$$

เมื่อ $X_1, \ldots, X_n$ เป็นตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบเบอร์นูลีโดยที่ $E(X_i) = p$ และ $V(X_i) = p(1-p)$

ดังนั้น จากทฤษฎีบทลิมิตเข้าสู่ส่วนกลาง เมื่อ $n \to \infty$ จะได้ว่าตัวแปรสุ่ม $\hat{p} = \frac{x}{n} = \frac{X_1 + \ldots + X_n}{n}$ มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ p และความแปรปรวนเท่ากับ $\frac{p(1-p)}{n}$ นั่นคือ

$$\hat{p} = \frac{x}{n} = \frac{X_1 + \ldots + X_n}{n} \sim N(p, \frac{p(1-p)}{n})$$

<u>หมายเหตุ</u>

จากทฤษฎีบทที่ 5.3.5 จะใช้ได้เมื่อ $n \to \infty$ แต่ในทางปฏิบัตินิยมใช้เมื่อ $n \geq 30$

5.4 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับค่าเฉลี่ยของกลุ่มตัวอย่าง

ในหัวข้อนี้จะกล่าวถึงการแจกแจงของตัวแปรสุ่มที่เกี่ยวข้องกับค่าเฉลี่ยของกลุ่มตัวอย่าง โดยอาศัยทฤษฎีบทที่ 5.3.2 ทฤษฎีบทลิมิตเข้าสู่ส่วนกลางในหัวข้อที่ 5.3 ทำให้สามารถสรุปเกี่ยวกับการแจกแจงของค่าเฉลี่ยของกลุ่มตัวอย่างกรณีที่กลุ่มตัวอย่างมีขนาดใหญ่หรือ $n \geq 30$ ได้ดังนี้

กรณีที่ 1 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับค่าเฉลี่ยของกลุ่มตัวอย่าง กรณีกลุ่มตัวอย่างมีขนาดใหญ่ ($n \geq 30$)

เมื่อประชากรมีการแจกแจงแบบใด ๆ โดยมีค่าเฉลี่ยเท่ากับ μ

(1) ถ้าทราบค่าความแปรปรวนของประชากร σ^2 แล้วค่าเฉลี่ยของกลุ่มตัวอย่างจะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

(2) ถ้าไม่ทราบค่าความแปรปรวนของประชากร σ^2 แล้วค่าเฉลี่ยของกลุ่มตัวอย่างจะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{s^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{s^2}{n})$$

__ตัวอย่างที่ 5.3.1__ ถ้าอายุการใช้งานของปรอทวัดไข้ชนิดหนึ่งมีค่าเฉลี่ยเท่ากับ 54 ครั้ง ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 6 ครั้ง ถ้าโรงพยาบาลแห่งหนึ่งเลือกซื้อปรอทวัดไข้อย่างสุ่มมาจำนวน 50 อัน จงหา

(1) ความน่าจะเป็นที่ปรอทวัดไข้ที่โรงพยาบาลแห่งนี้ซื้อมาจะมีอายุการใช้งานเฉลี่ยน้อยกว่า 52 ครั้ง

(2) ความน่าจะเป็นที่ปรอทวัดไข้ที่โรงพยาบาลแห่งนี้ซื้อมาจะมีอายุการใช้งานเฉลี่ยอยู่ระหว่าง 50 ถึง 60 ครั้ง

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนอายุการใช้งานของปรอทวัดไข้ จากโจทย์ $\mu = 54$, $\sigma = 6$ และ $n = 50$ ดังนั้น ค่าเฉลี่ยของอายุการใช้งานของปรอทวัดไข้ที่โรงพยาบาลแห่งนี้ซื้อมาทั้ง 50 อัน จะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \bar{X} \sim N(54, \frac{36}{50})$$

(1) ความน่าจะเป็นที่ปรอทวัดไข้ที่โรงพยาบาลแห่งนี้ซื้อมาจะมีอายุการใช้งานเฉลี่ยน้อยกว่า 52 ครั้ง เขียนแทนได้ด้วยสัญลักษณ์ $P(\bar{X} < 52)$ ซึ่งหาค่าได้ดังนี้

$$P(\bar{X} < 52) = P\left(\frac{\bar{X}-54}{\sqrt{36/50}} < \frac{52-54}{\sqrt{36/50}}\right)$$
$$= P(Z < -2.36)$$
$$= P(Z > 2.36)$$
$$= 1 - P(Z \le 2.36)$$
$$= 1 - 0.9909$$
$$= 0.0091$$

ดังนั้น ปรอทวัดไข้ที่โรงพยาบาลแห่งนี้ซื้อมามีอายุการใช้งานเฉลี่ยน้อยกว่า 52 ครั้ง ด้วยความน่าจะเป็นเท่ากับ 0.0091

(2) ความน่าจะเป็นที่ปรอทวัดไข้ที่โรงพยาบาลแห่งนี้ซื้อมาจะมีอายุการใช้งานเฉลี่ยอยู่ระหว่าง 50 ถึง 60 ครั้ง เขียนแทนได้ด้วยสัญลักษณ์ $P(50 < \bar{X} < 60)$ ซึ่งหาค่าได้ดังนี้

$$P(50 < \bar{X} < 60) = P\left(\frac{50-54}{\sqrt{36/50}} < \frac{\bar{X}-54}{\sqrt{36/50}} < \frac{60-54}{\sqrt{36/50}}\right)$$
$$= P(-4.71 < Z < 7.07)$$
$$= P(Z < 7.07) - P(Z \le -4.71)$$
$$= P(Z < 7.07) - P(Z \ge 4.71)$$
$$= P(Z < 7.07) - \{1 - P(Z < 4.71)\}$$
$$= 1 - \{1 - 1\}$$
$$= 1$$

ดังนั้น ปรอทวัดไข้ที่โรงพยาบาลแห่งนี้ซื้อมามีอายุการใช้งานเฉลี่ยอยู่ระหว่าง 50 ถึง 60 ครั้ง ด้วยความน่าจะเป็นเท่ากับ 1 ■

<u>**ตัวอย่างที่ 5.3.2**</u> ถ้าเงินเดือนของพนักงานบริษัทแห่งหนึ่งมีค่าเฉลี่ยเท่ากับ 15,000 บาท ถ้าสุ่มพนักงานจากบริษัทแห่งนี้มา 45 คน หาส่วนเบี่ยงเบนมาตรฐานของเงินเดือนพนักงานกลุ่มนี้ได้เท่ากับ 5,000 บาท จงหา

(1) ความน่าจะเป็นที่เงินเดือนเฉลี่ยของพนักงานที่สุ่มมามีค่ามากกว่า 16,000 บาท

(2) ความน่าจะเป็นที่เงินเดือนเฉลี่ยของพนักงานที่สุ่มมามีค่าตั้งแต่ 16,000 ถึง 17,000 บาท

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนเงินเดือนของพนักงานบริษัทแห่งนี้ จากโจทย์ $\mu = 15,000$, $n = 45$ และ $s = 5,000$ ดังนั้น เงินเดือนเฉลี่ยของพนักงานที่สุ่มมา 45 คน จะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{s^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{s^2}{n}) \Rightarrow \bar{X} \sim N(15000, \frac{25000000}{45})$$

(1) ความน่าจะเป็นที่เงินเดือนเฉลี่ยของพนักงานที่สุ่มมามีค่ามากกว่า 16,000 บาท เขียนแทนได้ด้วยสัญลักษณ์ $P(\bar{X} > 16,000)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(\bar{X} > 16,000) &= P\left(\frac{\bar{X}-15,000}{\sqrt{25,000,000/45}} < \frac{16,000-15,000}{\sqrt{25,000,000/45}}\right) \\ &= P(Z > 1.34) \\ &= 1 - P(Z \le 1.34) \\ &= 1 - 0.9099 \\ &= 0.0901 \end{aligned}$$

ดังนั้น เงินเดือนเฉลี่ยของพนักงานที่สุ่มมามีค่ามากกว่า 16,000 บาท ด้วยความน่าจะเป็นเท่ากับ 0.0901

(2) ความน่าจะเป็นที่เงินเดือนเฉลี่ยของพนักงานที่สุ่มมามีค่าตั้งแต่ 16,000 ถึง 17,000 บาท เขียนแทนได้ด้วยสัญลักษณ์ $P(16,000 \le \bar{X} \le 17,000)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(16,000 \le \bar{X} \le 17,000) &= P\left(\frac{16,000-15,000}{\sqrt{25,000,000/45}} \le \frac{\bar{X}-15,000}{\sqrt{25,000,000/45}} \le \frac{17,000-15,000}{\sqrt{25,000,000/45}}\right) \\ &= P(1.34 \le Z \le 2.68) \\ &= P(Z \le 2.68) - P(Z < 1.34) \\ &= 0.9963 - 0.9099 \\ &= 0.0864 \end{aligned}$$

ดังนั้น เงินเดือนเฉลี่ยของพนักงานที่สุ่มมามีค่าตั้งแต่ 16,000 ถึง 17,000 บาท ด้วยความน่าจะเป็นเท่ากับ 0.0864 ■

จากทฤษฎีบทที่ 5.3.1 และ 5.3.4 ในหัวข้อที่ 5.3 ทำให้สามารถสรุปเกี่ยวกับการแจกแจงของตัวแปรสุ่มที่เกี่ยวข้องกับค่าเฉลี่ยของกลุ่มตัวอย่างกรณีที่กลุ่มตัวอย่างมีขนาดเล็กหรือ $n < 30$ ได้ดังนี้

กรณีที่ 2 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับค่าเฉลี่ยของกลุ่มตัวอย่าง กรณีกลุ่มตัวอย่างมีขนาดเล็ก ($n < 30$)

เมื่อประชากรมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ

(1) ถ้าทราบค่าความแปรปรวนของประชากร σ^2 แล้วค่าเฉลี่ยของกลุ่มตัวอย่างจะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

(2) ถ้าไม่ทราบค่าความแปรปรวนของประชากร σ^2 แล้วตัวแปรสุ่ม $T = \frac{\bar{X}-\mu}{s/\sqrt{n}}$ จะมีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$T = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$$

<u>ตัวอย่างที่ 5.3.3</u> ถ้าน้ำหนักของคนไทยมีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเท่ากับ 55 กิโลกรัม และส่วนเบี่ยงเบนมาตรฐาน 18 กิโลกรัม

(1) ถ้าสุ่มเลือกคนไทยมา 12 คน จงหาความน่าจะเป็นที่น้ำหนักเฉลี่ยของคนกลุ่มนี้จะมีค่ามากกว่า 68 กิโลกรัม

(2) ถ้าสุ่มเลือกคนไทยมา 20 คน จงหาความน่าจะเป็นที่น้ำหนักเฉลี่ยของคนกลุ่มนี้จะมีค่ามากกว่า 68 กิโลกรัม

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนน้ำหนักของคนไทย จากโจทย์ $\mu = 55$ และ $\sigma = 18$

(1) เนื่องจากสุ่มเลือกคนไทยมา 12 คน ($n = 12$) ดังนั้น น้ำหนักเฉลี่ยของคนไทยที่สุ่มมา 12 คน มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \bar{X} \sim N(55, \frac{324}{12})$$

พิจารณาหาความน่าจะเป็นที่น้ำหนักเฉลี่ยของคนกลุ่มนี้จะมีค่ามากกว่า 68 กิโลกรัม เขียนแทนได้ด้วยสัญลักษณ์ $P(\bar{X} > 68)$

$$\begin{aligned} P(\bar{X} > 68) &= P\left(\frac{\bar{X} - 55}{\sqrt{324/12}} > \frac{68 - 55}{\sqrt{324/12}}\right) \\ &= P(Z > 2.50) \\ &= 1 - P(Z \le 2.50) \\ &= 1 - 0.9938 \\ &= 0.0062 \end{aligned}$$

ดังนั้น ถ้าสุ่มเลือกคนไทยมา 12 คน น้ำหนักเฉลี่ยของคนกลุ่มนี้จะมีค่ามากกว่า 68 กิโลกรัม ด้วยความน่าจะเป็นเท่ากับ 0.0062

(2) เนื่องจากสุ่มเลือกคนไทยมา 20 คน ($n = 20$) ดังนั้น น้ำหนักเฉลี่ยของคนไทยที่สุ่มมา 20 คน มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \bar{X} \sim N(55, \frac{324}{20})$$

ความน่าจะเป็นที่น้ำหนักเฉลี่ยของคนกลุ่มนี้จะมีค่ามากกว่า 68 กิโลกรัม เขียนแทนได้ด้วยสัญลักษณ์ $P(\bar{X} > 68)$

$$P(\bar{X} > 68) = P\left(\frac{\bar{X}-55}{\sqrt{324/20}} > \frac{68-55}{\sqrt{324/20}}\right)$$
$$= P(Z > 3.23)$$
$$= 1 - P(Z \le 3.23)$$
$$= 1 - 0.9994$$
$$= 0.0006$$

ดังนั้น ถ้าสุ่มเลือกคนไทยมา 20 คน น้ำหนักเฉลี่ยของคนกลุ่มนี้จะมีค่ามากกว่า 68 กิโลกรัม ด้วยความน่าจะเป็นเท่ากับ 0.0006 ■

ตัวอย่างที่ 5.3.4 ถ้าอายุการใช้งานของหลอดไฟชนิดหนึ่งมีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเท่ากับ 12,000 ชั่วโมง ถ้าสุ่มหลอดไฟมาทั้งหมด 9 หลอด เพื่อตรวจสอบอายุการใช้งาน แล้วคำนวณค่าส่วนเบี่ยงเบนมาตรฐานของอายุการใช้งานได้เท่ากับ 2,980 ชั่วโมง จงหา

(1) ความน่าจะเป็นที่อายุการใช้งานเฉลี่ยของหลอดไฟทั้ง 9 หลอด จะมีค่ามากกว่า 13,388 ชั่วโมง

(2) ความน่าจะเป็นที่อายุการใช้งานเฉลี่ยของหลอดไฟทั้ง 9 หลอด จะมีค่าตั้งแต่ 11,458 – 14,291 ชั่วโมง

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนอายุการใช้งานของหลอดไฟ จากโจทย์ $\mu = 12,000$, $n = 9$ และ $s = 2,980$ ดังนั้น

$$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1) \Rightarrow T = \frac{\bar{X}-12,000}{2,980/\sqrt{9}} \sim t(8)$$

(1) ความน่าจะเป็นที่อายุการใช้งานเฉลี่ยของหลอดไฟทั้ง 9 หลอด จะมีค่ามากกว่า 13,388 ชั่วโมง แทนได้ด้วยสัญลักษณ์ $P(\bar{X} > 13,388)$ ซึ่งหาค่าได้ดังนี้

$$P(\bar{X} > 13,388) = P\left(\frac{\bar{X}-12,000}{2,980/\sqrt{9}} > \frac{13,388-12,000}{2,980/\sqrt{9}}\right)$$
$$= P(T > 1.397)$$
$$= 1 - P(T \le 1.397)$$
$$= 1 - 0.90$$
$$= 0.10$$

ดังนั้น อายุการใช้งานเฉลี่ยของหลอดไฟทั้ง 9 หลอด จะมีค่ามากกว่า 13,388 ชั่วโมง ด้วยความน่าจะเป็นเท่ากับ 0.10

(2) ความน่าจะเป็นที่อายุการใช้งานเฉลี่ยของหลอดไฟทั้ง 9 หลอด จะมีค่าตั้งแต่ 11,458 – 14,291 ชั่วโมง แทนได้ด้วยสัญลักษณ์ $P(11,458 \le \bar{X} \le 14,291)$ ซึ่งหาค่าได้ดังนี้

$$P(11,458 \le \bar{X} \le 14,291) = P\left(\frac{11,458-12,000}{2,980/\sqrt{9}} \le \frac{\bar{X}-12,000}{2,980/\sqrt{9}} \le \frac{14,291-12,000}{2,980/\sqrt{9}}\right)$$
$$= P(-0.546 \le T \le 2.306)$$
$$= P(T \le 2.306) - P(T < -0.546)$$
$$= P(T \le 2.306) - P(T > 0.546)$$
$$= P(T \le 2.306) - \{1 - P(T \le 0.546)\}$$
$$= 0.975 - \{1 - 0.70\}$$
$$= 0.675$$

ดังนั้น อายุการใช้งานเฉลี่ยของหลอดไฟทั้ง 9 หลอด จะมีค่าตั้งแต่ 11,458 – 14,291 ชั่วโมง ด้วยความน่าจะเป็นเท่ากับ 0.675 ■

<u>ตัวอย่างที่ 5.3.5</u> ปริมาณวิตามินบีในอาหารเสริมชนิดหนึ่งมีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเท่ากับ 50 มิลลิกรัม ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 5 มิลลิกรัม จงหาขนาดของกลุ่มตัวอย่างเมื่อต้องการให้มีอย่างน้อย 99% ที่ปริมาณวิตามินบีเฉลี่ยของกลุ่มตัวอย่างมีค่าต่างจากปริมาณวิตามินบีเฉลี่ยของประชากรไม่เกิน 2 มิลลิกรัม

วิธีทำ กำหนดให้ X เป็นตัวแปรสุ่มแทนปริมาณวิตามินบีในอาหารเสริมชนิดนี้ จากโจทย์ $\mu = 50$ และ $\sigma = 5$ ดังนั้น ปริมาณวิตามินบีเฉลี่ยของกลุ่มตัวอย่างมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ $\frac{\sigma^2}{n}$ นั่นคือ

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \bar{X} \sim N(50, \frac{25}{n})$$

เนื่องจากโจทย์ต้องการให้มีอย่างน้อย 99% ที่ปริมาณวิตามินบีเฉลี่ยของกลุ่มตัวอย่างมีค่าต่างจากปริมาณวิตามินบีเฉลี่ยของประชากรไม่เกิน 2 มิลลิกรัม นั่นคือ ต้องการหาค่า n ที่ทำให้

$$P\left(\left|\bar{X} - \mu\right| \le 2\right) \ge 0.99$$

พิจารณา

$$P\left(\left|\bar{X} - \mu\right| \le 2\right) = P\left(\left|\bar{X} - 50\right| \le 2\right)$$
$$= P\left(-2 \le \bar{X} - 50 \le 2\right)$$
$$= P\left(\frac{-2}{\sqrt{25/n}} \le \frac{\bar{X}-50}{\sqrt{25/n}} \le \frac{2}{\sqrt{25/n}}\right)$$
$$= P\left(\frac{-2}{\sqrt{25/n}} \le Z \le \frac{2}{\sqrt{25/n}}\right)$$
$$= P\left(Z \le \frac{2}{\sqrt{25/n}}\right) - P\left(Z < \frac{-2}{\sqrt{25/n}}\right)$$

$$= P\left(Z \le \frac{2}{\sqrt{25/n}}\right) - P\left(Z > \frac{2}{\sqrt{25/n}}\right)$$

$$= P\left(Z \le \frac{2}{\sqrt{25/n}}\right) - \left\{1 - P\left(Z \le \frac{2}{\sqrt{25/n}}\right)\right\}$$

$$= 2P\left(Z \le \frac{2}{\sqrt{25/n}}\right) - 1$$

ดังนั้น

$$2P\left(Z \le \frac{2}{\sqrt{25/n}}\right) - 1 \ge 0.99$$

$$P\left(Z \le \frac{2}{\sqrt{25/n}}\right) \ge 0.995$$

เนื่องจาก $P(Z \le 2.576) = 0.995$ ดังนั้น ขนาดของกลุ่มตัวอย่าง n จะต้องสอดคล้องกับเงื่อนไข

$$\frac{2}{\sqrt{25/n}} \ge 2.576$$

$$\sqrt{n} \ge \frac{2.576\sqrt{25}}{2}$$

$$\sqrt{n} \ge 6.44$$

$$n \ge 41.4736$$

ดังนั้น ถ้าต้องการให้มีอย่างน้อย 99% ที่ปริมาณวิตามินบีเฉลี่ยของกลุ่มตัวอย่างมีค่าต่างจากปริมาณวิตามินบีเฉลี่ยของประชากรไม่เกิน 2 มิลลิกรัม จะต้องสุ่มตัวอย่างที่มีขนาดอย่างน้อยเท่ากับ 42 ■

5.5 การแจกแจงความน่าจะเป็นของค่าสัดส่วนของกลุ่มตัวอย่าง

ในหัวข้อนี้ได้กล่าวถึงการแจกแจงความน่าจะเป็นของค่าสัดส่วนของกลุ่มตัวอย่าง โดยอาศัยทฤษฎีบทที่ 5.3.5 ในหัวข้อที่ 5.3 ทำให้สามารถสรุปเกี่ยวกับการแจกแจงของค่าสัดส่วนของกลุ่มตัวอย่างได้ดังนี้

การแจกแจงความน่าจะเป็นของค่าสัดส่วนของกลุ่มตัวอย่าง

ถ้ากลุ่มตัวอย่างมีขนาดใหญ่ ($n \ge 30$) แล้วค่าสัดส่วนของกลุ่มตัวอย่างจะมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ p และความแปรปรวนเท่ากับ $\frac{p(1-p)}{n}$ นั่นคือ

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

<u>ตัวอย่างที่ 5.4.1</u> บริษัทผลิตแบตเตอรีแห่งหนึ่งทราบว่าโดยปกติแล้วจะมีแบตเตอรีที่ไม่ได้มาตรฐานปนมา 10% ของแบตเตอรีที่ผลิตทั้งหมด

(1) ถ้าสุ่มแบตเตอรีของบริษัทนี้มา 500 ลูก จงหาความน่าจะเป็นที่จะมีแบตเตอรีที่ไม่ได้มาตรฐานปนมามากกว่า 11%

(2) ถ้าสุ่มแบตเตอรีของบริษัทนี้มา 600 ลูก จงหาความน่าจะเป็นที่จะมีแบตเตอรีที่ไม่ได้มาตรฐานปนมามากกว่า 12%

วิธีทำ กำหนดให้ p แทนสัดส่วนของแบตเตอรีที่ไม่ได้มาตรฐานที่ผลิตจากบริษัทแห่งนี้ จากโจทย์ $p = 0.1$

(1) เนื่องจากสุ่มแบตเตอรีมาทั้งหมด 500 ลูก ($n = 500$) ดังนั้น สัดส่วนของแบตเตอรีที่ไม่ได้มาตรฐานจากแบตเตอรีที่สุ่มมา 500 ลูก มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ p และความแปรปรวนเท่ากับ $\frac{p(1-p)}{n}$ นั่นคือ

$$\hat{p} \sim N(p, \frac{p(1-p)}{n}) \Rightarrow \hat{p} \sim N(0.1, 0.00018)$$

ถ้าสุ่มแบตเตอรีของบริษัทนี้มา 500 ลูก ความน่าจะเป็นที่จะมีแบตเตอรีที่ไม่ได้มาตรฐานปนมามากกว่า 11% แทนได้ด้วยสัญลักษณ์ $P(\hat{p} > 0.11)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(\hat{p} > 0.11) &= P\left(\frac{\hat{p}-0.1}{\sqrt{0.00018}} > \frac{0.11-0.1}{\sqrt{0.00018}}\right) \\ &= P(Z > 0.75) \\ &= 1 - P(Z \leq 0.75) \\ &= 1 - 0.7734 \\ &= 0.2266 \end{aligned}$$

ดังนั้น ถ้าสุ่มตัวอย่างแบตเตอรีของบริษัทนี้มา 500 ลูก จะพบแบตเตอรีที่ไม่ได้มาตรฐานปนมามากกว่า 11% ด้วยความน่าจะเป็นเท่ากับ 0.2266

(2) เนื่องจากสุ่มแบตเตอรีมาทั้งหมด 600 ลูก ($n = 600$) ดังนั้น สัดส่วนของแบตเตอรีที่ไม่ได้มาตรฐานจากแบตเตอรีที่สุ่มมา 600 ลูก มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ p และความแปรปรวนเท่ากับ $\frac{p(1-p)}{n}$ นั่นคือ

$$\hat{p} \sim N(p, \frac{p(1-p)}{n}) \Rightarrow \hat{p} \sim N(0.1, 0.00015)$$

ถ้าสุ่มแบตเตอรีของบริษัทนี้มา 600 ลูก ความน่าจะเป็นที่จะมีแบตเตอรีที่ไม่ได้มาตรฐานปนมามากกว่า 12% แทนได้ด้วยสัญลักษณ์ $P(\hat{p} > 0.12)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(\hat{p} > 0.12) &= P\left(\frac{\hat{p}-0.1}{\sqrt{0.00015}} > \frac{0.12-0.1}{\sqrt{0.00015}}\right) \\ &= P(Z > 1.63) \\ &= 1 - P(Z \leq 1.63) \\ &= 1 - 0.9484 \end{aligned}$$

$$= 0.0516$$

ดังนั้น ถ้าสุ่มตัวอย่างแบตเตอรีของบริษัทนี้มา 600 ลูก จะพบแบตเตอรีที่ไม่ได้มาตรฐานปนมามากกว่า 12% ด้วยความน่าจะเป็นเท่ากับ 0.0516 ■

<u>**ตัวอย่างที่**</u> <u>5.4.2</u> จากการสำรวจพนักงานบริษัทหนึ่งพบว่า พนักงานชายที่รับประทานอาหารเช้าคิดเป็น 5% ของพนักงานชายทั้งหมด และพนักงานหญิงที่รับประทานอาหารเช้าคิดเป็น 12% ของพนักงานหญิงทั้งหมด

(1) ถ้าสุ่มพนักงานชายของบริษัทนี้มาจำนวน 100 คน จงหาความน่าจะเป็นที่จะมีพนักงานชายที่รับประทานอาหารเช้าปนมาตั้งแต่ 6% - 7%

(2) ถ้าสุ่มพนักงานหญิงของบริษัทนี้มาจำนวน 200 คน จงหาความน่าจะเป็นที่จะมีพนักงานหญิงที่รับประทานอาหารเช้าปนมาน้อยกว่า 10%

วิธีทำ (1) กำหนดให้ p แทนสัดส่วนของพนักงานชายที่รับประทานอาหารเช้าเมื่อเปรียบเทียบกับพนักงานชายทั้งหมดของบริษัทแห่งนี้ จากโจทย์ $p = 0.05$ เนื่องจากสุ่มพนักงานชายมาทั้งหมด 100 คน ($n = 100$) ดังนั้นสัดส่วนของพนักงานชายที่รับประทานอาหารเช้าเมื่อเปรียบเทียบกับพนักงานชายทั้งหมดที่สุ่ม 100 คน มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ p และความแปรปรวนเท่ากับ $\frac{p(1-p)}{n}$ นั่นคือ

$$\hat{p} \sim N(p, \frac{p(1-p)}{n}) \Rightarrow \hat{p} \sim N(0.05, 0.000475)$$

ถ้าสุ่มพนักงานชายของบริษัทนี้มาจำนวน 100 คน ความน่าจะเป็นที่จะมีพนักงานชายที่รับประทานอาหารเช้าปนมาตั้งแต่ 6% - 7% แทนได้ด้วยสัญลักษณ์ $P(0.06 \le \hat{p} \le 0.07)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(0.06 \le \hat{p} \le 0.07) &= P\left(\frac{0.06-0.05}{\sqrt{0.000475}} \le \frac{\hat{p}-0.05}{\sqrt{0.000475}} \le \frac{0.07-0.05}{\sqrt{0.000475}}\right) \\ &= P(0.46 \le Z \le 0.92) \\ &= P(Z \le 0.92) - P(Z < 0.46) \\ &= 0.8212 - 0.6772 \\ &= 0.1440 \end{aligned}$$

ดังนั้น ถ้าสุ่มพนักงานชายของบริษัทนี้มาจำนวน 100 คน จะมีพนักงานชายที่รับประทานอาหารเช้าปนมาตั้งแต่ 6% - 7% ด้วยความน่าจะเป็นเท่ากับ 0.1440

(2) กำหนดให้ q แทนสัดส่วนของพนักงานหญิงที่รับประทานอาหารเช้าเมื่อเปรียบเทียบกับพนักงานหญิงทั้งหมดของบริษัทแห่งนี้ จากโจทย์ $q = 0.12$ เนื่องจากสุ่มพนักงานหญิงมาทั้งหมด 200 คน ($n = 200$) ดังนั้นสัดส่วนของพนักงานหญิงที่รับประทานอาหารเช้าเมื่อเปรียบเทียบกับพนักงานหญิงทั้งหมดที่สุ่ม 200 คน มีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ q และความแปรปรวนเท่ากับ $\frac{q(1-q)}{n}$ นั่นคือ

$$\hat{q} \sim N(q, \frac{q(1-q)}{n}) \Rightarrow \hat{q} \sim N(0.12, 0.000528)$$

ถ้าสุ่มพนักงานหญิงของบริษัทนี้มาจำนวน 200 คน ความน่าจะเป็นที่จะมีพนักงานหญิงที่รับประทานอาหารเช้าปนมาน้อยกว่า 10% แทนได้ด้วยสัญลักษณ์ $P(\hat{q} < 0.10)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(\hat{q} < 0.10) &= P\left(\frac{\hat{q}-0.12}{\sqrt{0.000528}} < \frac{0.10-0.12}{\sqrt{0.000528}}\right) \\ &= P(Z < -0.87) \\ &= P(Z > 0.87) \\ &= 1 - P(Z \le 0.87) \\ &= 1 - 0.8078 \\ &= 0.1922 \end{aligned}$$

ดังนั้น ถ้าสุ่มพนักงานหญิงของบริษัทนี้มาจำนวน 200 คน จะมีพนักงานหญิงที่รับประทานอาหารเช้าปนมาน้อยกว่า 10% ด้วยความน่าจะเป็นเท่ากับ 0.1922 ■

5.6 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

ในหัวข้อนี้ได้กล่าวถึงการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง โดยทฤษฎีบทที่ 5.3.3 ในหัวข้อที่ 5.3 ทำให้สามารถสรุปเกี่ยวกับการแจกแจงของตัวแปรสุ่มที่เกี่ยวข้องกับความแปรปรวน (หรือส่วนเบี่ยงเบนมาตรฐาน) ของกลุ่มตัวอย่างได้ดังนี้

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

ถ้าประชากรมีการแจกแจงแบบปกติ แล้วตัวแปรสุ่ม $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ จะมีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

<u>ตัวอย่างที่ 5.5.1</u> พิจารณาประชากรที่มีการแจกแจงแบบปกติซึ่งมีความแปรปรวนของประชากรเท่ากับ 6 ถ้าสุ่มตัวอย่างขนาด 25 จงหา

(1) ความน่าจะเป็นที่ความแปรปรวนของกลุ่มตัวอย่างมีค่ามากกว่า 9.10375

(2) ความน่าจะเป็นที่ความแปรปรวนของกลุ่มตัวอย่างมีค่าอยู่ระหว่าง 3.462 และ 10.745

(3) ความน่าจะเป็นที่ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างมีค่าน้อยกว่า 1.76075

วิธีทำ จากโจทย์ $\sigma^2 = 6$ และ $n = 25$ ดังนั้น $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1) \Rightarrow \chi^2 = \frac{24s^2}{6} \sim \chi^2(24)$$

(1) ความน่าจะเป็นที่ความแปรปรวนของกลุ่มตัวอย่างมีค่ามากกว่า 9.10375 เขียนแทนได้ด้วยสัญลักษณ์ $P(s^2 > 9.10375)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(s^2 > 9.10375) &= P\left(\frac{24s^2}{6} > \frac{24(9.10375)}{6}\right) \\ &= P(\chi^2 > 36.415) \\ &= 1 - P(\chi^2 \leq 36.415) \\ &= 1 - 0.95 \\ &= 0.05 \end{aligned}$$

ดังนั้น ความแปรปรวนของกลุ่มตัวอย่างมีค่ามากกว่า 9.10375 ด้วยความน่าจะเป็นเท่ากับ 0.05

(2) ความน่าจะเป็นที่ความแปรปรวนของกลุ่มตัวอย่างมีค่าอยู่ระหว่าง 3.462 และ 10.745 เขียนแทนได้ด้วยสัญลักษณ์ $P(3.462 < s^2 < 10.745)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(3.462 < s^2 < 10.745) &= P\left(\frac{24(3.462)}{6} < \frac{24s^2}{6} > \frac{24(10.745)}{6}\right) \\ &= P(13.848 < \chi^2 < 42.98) \\ &= P(\chi^2 < 42.98) - P(\chi^2 \leq 13.848) \\ &= 0.99 - 0.05 \\ &= 0.94 \end{aligned}$$

ดังนั้น ความแปรปรวนของกลุ่มตัวอย่างมีค่าอยู่ระหว่าง 3.462 และ 10.745 ด้วยความน่าจะเป็นเท่ากับ 0.94

(3) ความน่าจะเป็นที่ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างมีค่าน้อยกว่า 1.76075 เขียนแทนได้ด้วยสัญลักษณ์ $P(s < 1.76075)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(s < 1.76075) &= P\left(\frac{24s^2}{6} < \frac{24(1.76075)^2}{6}\right) \\ &= P(\chi^2 < 12.401) \\ &= 0.025 \end{aligned}$$

ดังนั้น ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างมีค่าน้อยกว่า 1.76075 ด้วยความน่าจะเป็นเท่ากับ 0.025 ■

ตัวอย่างที่ 5.5.2 เวลาที่ใช้ในการทำงานชิ้นหนึ่งมีการแจกแจงแบบปกติด้วยส่วนเบี่ยงเบนมาตรฐานเท่ากับ 5 ถ้าสุ่มการทำงานชิ้นนี้มาทั้งหมด 15 ครั้ง จงหา

(1) ความน่าจะเป็นที่ความแปรปรวนของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่าไม่เกิน 46.641

(2) ความน่าจะเป็นที่ความแปรปรวนของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่าอยู่ระหว่าง 10.0518 และ 32.4125

(3) ความน่าจะเป็นที่ส่วนเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่ามากกว่า 4.6441

วิธีทำ จากโจทย์ $\sigma = 5$ และ $n = 15$ ดังนั้น $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$ มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ด้วยองศาอิสระเท่ากับ $n-1$ นั่นคือ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1) \Rightarrow \chi^2 = \frac{14s^2}{25} \sim \chi^2(14)$$

(1) ความน่าจะเป็นที่ความแปรปรวนของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่าไม่เกิน 46.641 เขียนแทนได้ด้วยสัญลักษณ์ $P(s^2 \le 46.641)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(s^2 \le 46.641) &= P\left(\frac{14s^2}{25} > \frac{14(46.641)}{25}\right) \\ &= P(\chi^2 \le 26.119) \\ &= 0.975 \end{aligned}$$

ดังนั้น ความแปรปรวนของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่าไม่เกิน 46.641 ด้วยความน่าจะเป็นเท่ากับ 0.975

(2) ความน่าจะเป็นที่ความแปรปรวนของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่าอยู่ระหว่าง 10.0518 และ 32.4125 เขียนแทนได้ด้วยสัญลักษณ์ $P(10.0518 < s^2 < 32.4125)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(10.0518 < s^2 < 32.4125) &= P\left(\frac{14(10.0518)}{25} < \frac{14s^2}{25} > \frac{14(32.4125)}{25}\right) \\ &= P(5.629 < \chi^2 < 18.151) \\ &= P(\chi^2 < 18.151) - P(\chi^2 \le 5.629) \\ &= 0.775 \end{aligned}$$

ดังนั้น ความแปรปรวนของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่าอยู่ระหว่าง 10.0518 และ 32.4125 ด้วยความน่าจะเป็นเท่ากับ 0.775

(3) ความน่าจะเป็นที่ส่วนเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่ามากกว่า 4.6441 เขียนแทนได้ด้วยสัญลักษณ์ $P(s > 4.6441)$ ซึ่งหาค่าได้ดังนี้

$$\begin{aligned} P(s > 4.6441) &= P\left(\frac{14s^2}{25} > \frac{14(4.6441)^2}{25}\right) \\ &= P(\chi^2 > 12.078) \\ &= 1 - P(\chi^2 \le 12.078) \\ &= 1 - 0.4 \end{aligned}$$

$$= 0.6$$

ดังนั้น ส่วนเบี่ยงเบนมาตรฐานของเวลาที่ใช้ในการทำงานชิ้นนี้มีค่ามากกว่า 4.6441 ด้วยความน่าจะเป็นเท่ากับ 0.6 ■

<u>หมายเหตุ</u>

การแจกแจงของฟังก์ชันที่ได้จากกลุ่มตัวอย่างนอกเหนือจากที่ได้กล่าวมาแล้วในบทนี้ยังมีอีกหลายฟังก์ชัน เช่น การแจกแจงของผลต่างระหว่างค่าเฉลี่ยของกลุ่มตัวอย่าง 2 กลุ่ม การแจกแจงของผลต่างระหว่างค่าสัดส่วนของกลุ่มตัวอย่าง 2 กลุ่ม และการแจกแจงของอัตราส่วนความแปรปรวนของกลุ่มตัวอย่าง 2 กลุ่ม เป็นต้น ซึ่งผู้อ่านสามารถหาอ่านเพิ่มเติมได้ในหนังสือเกี่ยวกับความน่าจะเป็นและสถิติ

บทที่ 6
การประมาณค่าพารามิเตอร์
(Parameter Estimation)

ในบทที่ 6 และ 7 จะกล่าวถึงการอนุมานเชิงสถิติ (statistical inference) ซึ่งเป็นกระบวนการที่ใช้ข้อมูลที่ได้จากกลุ่มตัวอย่าง เพื่อนำมาหาข้อสรุปเกี่ยวกับข้อมูลของประชากร การอนุมานเชิงสถิติแบ่งออกเป็น 2 ส่วน คือ การประมาณค่าพารามิเตอร์ (parameter estimation) และการทดสอบสมมติฐาน (test of hypothesis) สำหรับในบทนี้จะกล่าวถึงเฉพาะการประมาณค่าพารามิเตอร์ ประกอบไปด้วย การประมาณค่าเฉลี่ยของประชากร การประมาณค่าสัดส่วนของประชากร และการประมาณความแปรปรวนของประชากร (หรือส่วนเบี่ยงเบนมาตรฐานของประชากร) ซึ่งในบทนี้จะใช้ความรู้เกี่ยวกับการแจกแจงของฟังก์ชันที่ได้จากกลุ่มตัวอย่างที่ได้กล่าวมาแล้วในบทที่ 5 เป็นพื้นฐานสำหรับการประมาณค่าพารามิเตอร์ต่าง ๆ เหล่านี้ ส่วนการทดสอบสมมติฐานจะกล่าวถึงอีกครั้งในบทที่ 7 ต่อไป

6.1 การประมาณค่าพารามิเตอร์

การประมาณค่าพารามิเตอร์แบ่งออกเป็น 2 แบบ คือ

1. การประมาณค่าแบบจุด (point estimation) เป็นการประมาณค่าพารามิเตอร์ด้วยค่าสถิติที่ได้จากกลุ่มตัวอย่าง โดยที่ค่าดังกล่าวจะมีค่าเป็นค่าใดค่าหนึ่ง

เราอาจพบว่ามีค่าสถิติหลายค่าที่สามารถใช้ในการประมาณค่าพารามิเตอร์ได้ ดังนั้น เราจะเลือกค่าสถิติที่เป็นตัวประมาณค่าที่ไม่เอนเอียงซึ่งมีนิยามดังนี้

นิยามที่ 6.1.1 ค่าสถิติ $\hat{\theta}$ จะเรียกว่า ***ตัวประมาณค่าที่ไม่เอนเอียง (unbiased estimator)*** ของพารามิเตอร์ θ เมื่อค่าเฉลี่ยของ $\hat{\theta}$ มีค่าเท่ากับ θ นั่นคือ

$$E(\hat{\theta}) = \theta$$

บางครั้งตัวประมาณค่าที่ไม่เอนเอียงอาจมีหลายตัว ดังนั้น เราจะเลือกค่าสถิติที่เป็นตัวประมาณค่าที่มีประสิทธิภาพที่สุดซึ่งมีนิยามดังนี้

นิยามที่ 6.1.2 สำหรับตัวประมาณค่าที่ไม่เอนเอียง (unbiased estimator) ทั้งหมดของพารามิเตอร์ θ จะเรียก $\hat{\theta}_m$ ว่าเป็น ***ตัวประมาณค่าที่มีประสิทธิภาพที่สุด (the most efficient estimator)*** ของ θ ถ้าความแปรปรวนของ θ_m มีค่าน้อยที่สุดเมื่อเทียบกับตัวประมาณค่าอื่น

นั่นคือ ถ้าค่าสถิติ $\hat{\theta}_1$ และ $\hat{\theta}_2$ เป็นตัวประมาณค่าที่ไม่เอนเอียง (unbiased estimator) ของพารามิเตอร์ θ โดยที่ $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ แล้ว $\hat{\theta}_1$ เป็นตัวประมาณค่าที่มีประสิทธิภาพมากกว่า $\hat{\theta}_2$ ดังนั้น เราจะเลือก $\hat{\theta}_1$ เป็นค่าประมาณของพารามิเตอร์ θ

2. การประมาณค่าแบบช่วง (interval estimation) เป็นการประมาณค่าพารามิเตอร์โดยค่าที่ประมาณได้จะมีลักษณะเป็นช่วง นิยามที่สำคัญเกี่ยวกับการประมาณค่าแบบช่วงมีดังนี้

นิยามที่ 6.1.3

ช่วงความเชื่อมั่น (confidence interval) ของค่าพารามิเตอร์ θ หมายถึง ช่วงของการประมาณพารามิเตอร์ θ ซึ่งสามารถเขียนแทนช่วงความเชื่อมั่นได้ดังนี้

$$\hat{\theta}_L < \theta < \hat{\theta}_U$$

เมื่อ $\hat{\theta}_L$ และ $\hat{\theta}_U$ แทน ขอบเขตล่างและขอบเขตบนของช่วงความเชื่อมั่น ตามลำดับ

ระดับความเชื่อมั่น (confidence level) หมายถึง ความน่าจะเป็นที่พารามิเตอร์ θ จะอยู่ในช่วงความเชื่อมั่นที่ประมาณได้ ซึ่งจะแทนด้วยสัญลักษณ์ $1-\alpha$

<u>ข้อสังเกต</u> จากนิยามที่ 6.1.3

1. ช่วงความเชื่อมั่นมีความสัมพันธ์กับระดับความเชื่อมั่นดังนี้

$$P(\hat{\theta}_L < \theta < \hat{\theta}_U) = 1-\alpha$$

ดังนั้น ถ้าระดับความเชื่อมั่นเปลี่ยนไปจะทำให้ช่วงความเชื่อมั่นเปลี่ยนแปลงไปด้วย

2. เนื่องจากระดับความเชื่อมั่นหมายถึงความน่าจะเป็น ดังนั้น $0 \le 1-\alpha \le 1$ และ $0 \le \alpha \le 1$

6.2 การประมาณค่าเฉลี่ยของประชากร

6.2.1 การประมาณค่าเฉลี่ยของประชากรแบบจุด

ประมาณค่าเฉลี่ยประชากร (μ) แบบจุดด้วยค่าเฉลี่ยของกลุ่มตัวอย่าง ($\bar{X}$) โดยที่

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

เมื่อ X_i หมายถึง ค่าของสิ่งที่สนใจ
N หมายถึง ขนาดประชากร
n หมายถึง ขนาดกลุ่มตัวอย่าง
μ หมายถึง ค่าเฉลี่ยของประชากร
$\bar{X}$ หมายถึง ค่าเฉลี่ยของกลุ่มตัวอย่าง

พิจารณาการหาค่าเฉลี่ยของค่าเฉลี่ยของกลุ่มตัวอย่าง ($\bar{X}$) ดังนี้

$$E(\bar{X}) = E\left(\frac{X_1 + \ldots + X_n}{n}\right)$$
$$= \frac{1}{n}E\left(X_1 + \ldots + X_n\right)$$

จากสมบัติของค่าเฉลี่ยหรือค่าคาดหมาย จะได้ว่า

$$E(\bar{X}) = \frac{1}{n}\left(E(X_1) + \ldots + E(X_n)\right)$$

เนื่องจาก X_i เป็นตัวแปรสุ่มที่มีค่าเฉลี่ยเท่ากับ μ จึงได้ว่า

$$E(\bar{X}) = \frac{1}{n}(\mu + \ldots + \mu)$$
$$= \frac{n\mu}{n}$$
$$= \mu$$

ดังนั้น

$$E(\bar{X}) = \mu$$

เนื่องจาก $E(\bar{X}) = \mu$ จึงทำให้ $\bar{X}$ เป็นตัวประมาณค่าที่ไม่เอนเอียงของ μ ดังนั้น สำหรับการประมาณค่าเฉลี่ยของประชากร (μ) แบบจุด จึงสามารถประมาณได้ด้วยค่าเฉลี่ยของกลุ่มตัวอย่าง ($\bar{X}$) นั่นเอง

ตัวอย่างที่ 6.2.1 มีคะแนนสอบของนักศึกษา 600 คน สุ่มคะแนนนักศึกษามา 10 คน พบว่ามีคะแนนดังนี้

36, 42, 34, 40, 64, 50, 22, 54, 60, 32

จงประมาณค่าเฉลี่ยของคะแนนสอบนักศึกษา 600 คน โดยใช้การประมาณค่าแบบจุด

วิธีทำ การประมาณค่าแบบจุดสำหรับค่าเฉลี่ยของประชากร ทำได้โดยหาค่าเฉลี่ยของกลุ่มตัวอย่างดังนี้

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{10} X_i}{10} = \frac{36 + 42 + \ldots + 32}{10} = \frac{434}{10} = 43.4$$

ดังนั้นค่าเฉลี่ยของประชากรหรือค่าเฉลี่ยคะแนนสอบนักศึกษา 600 คน มีค่าประมาณ 43.4 คะแนน ■

6.2.2 การประมาณค่าเฉลี่ยของประชากรแบบช่วง

<u>กรณีที่ 1</u>: กลุ่มตัวอย่างมีขนาดใหญ่ ($n \geq 30$)

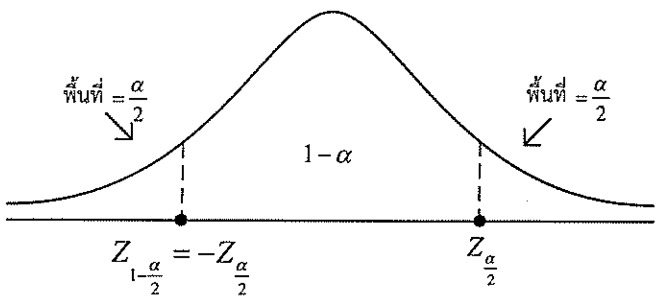
พิจารณาการแจกแจงความน่าจะเป็นสำหรับค่าเฉลี่ยของกลุ่มตัวอย่างขนาดใหญ่ เมื่อประชากรมีการแจกแจงแบบใด ๆ ด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 จากบทที่ 5 หัวข้อที่ 5.4 จะได้ว่า

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

นั่นคือ

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

พิจารณาที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ และกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (โค้งปกติมาตรฐาน) ดังรูป



นั่นคือ

$$P(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < Z_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \bar{X}-\mu < Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

พิจารณา

(1) กรณีที่ $\bar{X} - \mu > -Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ ดังนั้น

$$\bar{X} - \mu > -Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$-\mu > -\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$\mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

(2) กรณีที่ $\bar{X} - \mu < Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ ดังนั้น

$$\bar{X} - \mu < Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$-\mu < -\overline{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$\mu > \overline{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

จากทั้งสองกรณีสรุปได้ว่า

$$P(\overline{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

ดังนั้น ที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ ช่วงการประมาณค่าของ μ คือ

$$\overline{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

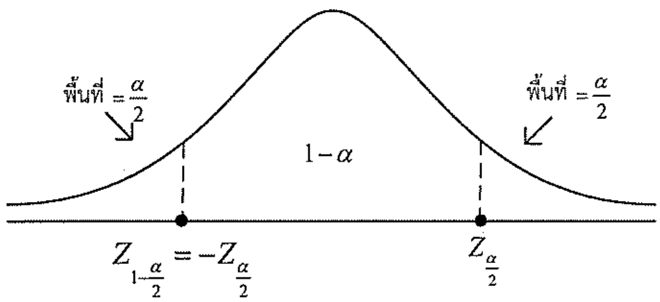
พิจารณาการแจกแจงความน่าจะเป็นสำหรับค่าเฉลี่ยของกลุ่มตัวอย่างขนาดใหญ่ เมื่อประชากรมีการแจกแจงแบบใด ๆ ด้วยค่าเฉลี่ยเท่ากับ μ แต่ไม่ทราบค่าความแปรปรวนของประชากร จากบทที่ 5 หัวข้อที่ 5.4 จะได้ว่า

$$\overline{X} \sim N(\mu, \frac{s^2}{n})$$

นั่นคือ

$$Z = \frac{\overline{X} - \mu}{s/\sqrt{n}} \sim N(0,1)$$

พิจารณาที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ และกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (โค้งปกติมาตรฐาน) ดังรูป



นั่นคือ

$$P(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}} < \frac{\overline{X} - \mu}{s/\sqrt{n}} < Z_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \overline{X} - \mu < Z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}) = 1-\alpha$$

พิจารณา

(1) กรณีที่ $\overline{X}-\mu > -Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$ ดังนั้น

$$\overline{X}-\mu > -Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$-\mu > -\overline{X}-Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$\mu < \overline{X}+Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

(2) กรณีที่ $\overline{X}-\mu < Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$ ดังนั้น

$$\overline{X}-\mu < Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$-\mu < -\overline{X}+Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$\mu > \overline{X}-Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

จากทั้งสองกรณีสรุปได้ว่า

$$P(\overline{X}-Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \overline{X}+Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}) = 1-\alpha$$

ดังนั้น ที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ ช่วงการประมาณค่าของ μ คือ

$$\overline{X}-Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \overline{X}+Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

การประมาณค่าแบบช่วงหรือการสร้างช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยของประชากร กรณีกลุ่มตัวอย่างมีขนาดใหญ่ ($n \geq 30$) สามารถสรุปได้ดังนี้

กรณีที่ 1 การประมาณค่าเฉลี่ยของประชากรแบบช่วงกรณีกลุ่มตัวอย่างมีขนาดใหญ่ ($n \geq 30$)

เมื่อประชากรมีการแจกแจงแบบใด ๆ โดยมีค่าเฉลี่ยเท่ากับ μ

(1) ถ้าทราบค่าความแปรปรวนของประชากร σ^2 ดังนั้น ช่วงการประมาณค่าของ μ ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$\overline{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

(2) ถ้าไม่ทราบค่าความแปรปรวนของประชากร σ^2 ดังนั้น ช่วงการประมาณค่าของ μ ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$\overline{X}-Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \overline{X}+Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

เมื่อ μ หมายถึง ค่าเฉลี่ยของประชากร

$\bar{X}$ หมายถึง ค่าเฉลี่ยของกลุ่มตัวอย่าง

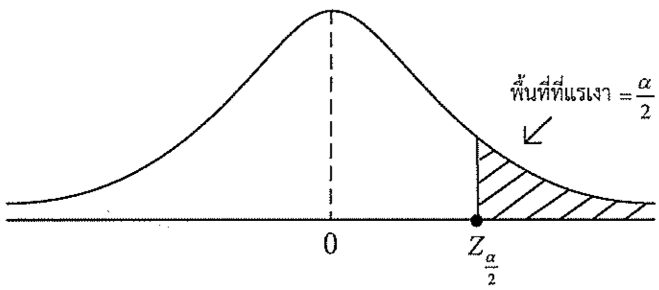
σ หมายถึง ส่วนเบี่ยงเบนมาตรฐานของประชากร

s หมายถึง ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

n หมายถึง ขนาดของกลุ่มตัวอย่าง

$1-\alpha$ หมายถึง ระดับความเชื่อมั่น

$Z_{\frac{\alpha}{2}}$ หมายถึง ค่าของตัวแปรสุ่ม $Z \sim N(0,1)$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน)

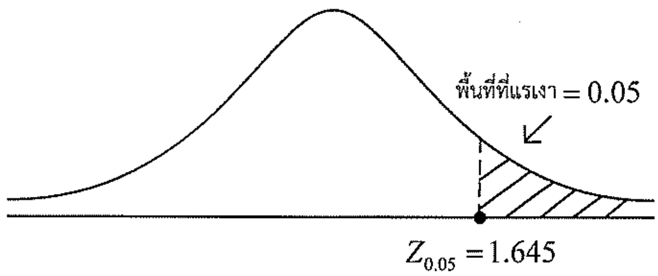


<u>ตัวอย่างที่ 6.2.2</u> โรงงานผลิตสายเบรกแห่งหนึ่งต้องการทราบว่า สายเบรกที่ผลิตได้จะสามารถรับแรงดึงได้โดยเฉลี่ยเท่าใด เพื่อที่จะประมาณค่าเฉลี่ยของแรงดึงนั้น ได้มีการเลือกตัวอย่างเชิงสุ่มมา 32 เส้น แล้วทำการตรวจสอบแรงดึงของแต่ละเส้นต่อจากนั้นจึงหาค่าเฉลี่ยของตัวอย่างได้ 42,196 ปอนด์ จากประสบการณ์ทางโรงงานทราบว่าส่วนเบี่ยงเบนมาตรฐานของแรงดึงมีค่าเท่ากับ 500 ปอนด์ จงประมาณค่าแรงดึงเฉลี่ยของสายเบรกที่ผลิตจากโรงงานนี้ที่ระดับความเชื่อมั่น 90%

วิธีทำ กำหนดให้ X แทนแรงดึงของสายเบรกที่ผลิตจากโรงงานแห่งนี้ จากโจทย์ ความแปรปรวนของประชากร ค่าเฉลี่ยของกลุ่มตัวอย่าง ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 500, 42196, 32 และ 0.90 ตามลำดับ นั่นคือ $\sigma = 500$, $\bar{X} = 42,196$, $n = 32$ และ $\alpha = 1-0.90 = 0.10$ ดังนั้น ค่าแรงดึงเฉลี่ยของสายเบรกที่ผลิตจากโรงงานแห่งนี้มีค่าอยู่ในช่วง

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ดังรูป



โดยใช้ตารางที่ 3.2 เมื่อ $F(z) = 1 - 0.05 = 0.95$ จะได้ว่า

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = Z_{0.05} = 1.645$$

ดังนั้น

$$42,196 - 1.645\frac{500}{\sqrt{32}} < \mu < 42,196 + 1.645\frac{500}{\sqrt{32}}$$

$$42,051 < \mu < 42,341$$

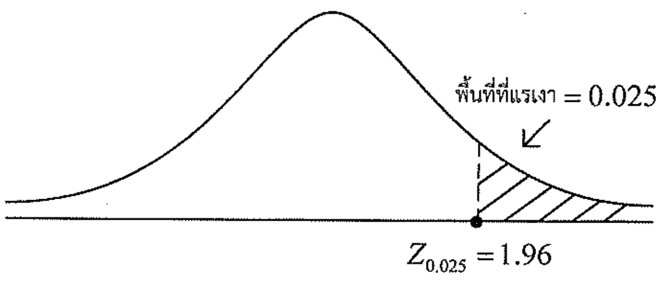
ดังนั้น แรงดึงเฉลี่ยของสายเบรกที่ผลิตจากโรงงานแห่งนี้มีค่าประมาณ 42,051 - 42,321 ปอนด์ ที่ระดับความเชื่อมั่น 90% ■

<u>ตัวอย่างที่ 6.2.3</u> นักวิทยาศาสตร์คนหนึ่งต้องการศึกษาเกี่ยวกับปริมาณออกซิเจนที่ละลายในแม่น้ำเจ้าพระยา จึงทำการเก็บตัวอย่างน้ำในแม่น้ำเจ้าพระยามาทั้งหมด 36 ตัวอย่าง แล้วนำมาคำนวณหาปริมาณออกซิเจนเฉลี่ยที่ละลายในน้ำได้เท่ากับ 5 มิลลิกรัมต่อลิตร และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 4 จงหาช่วงความเชื่อมั่นของปริมาณออกซิเจนเฉลี่ยที่ละลายในแม่น้ำเจ้าพระยาที่ระดับความเชื่อมั่น 95%

วิธีทำ กำหนดให้ X แทนปริมาณออกซิเจนที่ละลายในน้ำ จากโจทย์ ค่าเฉลี่ยของกลุ่มตัวอย่าง ความแปรปรวนของกลุ่มตัวอย่าง ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 5, 4, 36 และ 0.95 ตามลำดับ นั่นคือ $\bar{X} = 5$, $s = 4$, $n = 36$ และ $\alpha = 1 - 0.95 = 0.05$ ดังนั้น ปริมาณออกซิเจนเฉลี่ยที่ละลายในแม่น้ำเจ้าพระยามีค่าอยู่ในช่วง

$$\bar{X} - Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ดังรูป



โดยใช้ตารางที่ 3.2 เมื่อ $F(z)=1-0.025=0.975$ จะได้ว่า

$$Z_{\frac{\alpha}{2}}=Z_{\frac{0.05}{2}}=Z_{0.025}=1.96$$

ดังนั้น

$$5-1.96\frac{4}{\sqrt{36}}<\mu<5+1.96\frac{4}{\sqrt{36}}$$

$$3.6933<\mu<6.3067$$

ดังนั้น ปริมาณออกซิเจนเฉลี่ยที่ละลายในแม่น้ำเจ้าพระยามีค่าประมาณ 3.6933 – 6.3067 มิลลิกรัมต่อลิตร ที่ระดับความเชื่อมั่น 95% ■

<u>กรณีที่ 2</u>: กลุ่มตัวอย่างมีขนาดเล็ก ($n<30$)

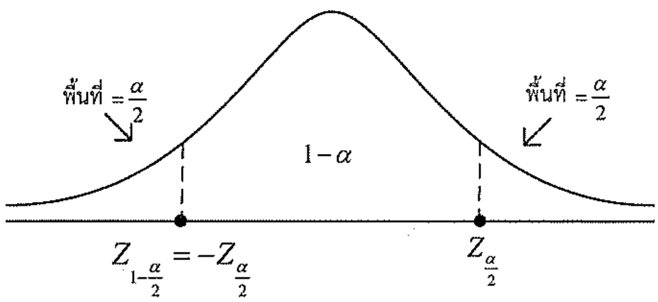
พิจารณาการแจกแจงความน่าจะเป็นสำหรับค่าเฉลี่ยของกลุ่มตัวอย่าง เมื่อกลุ่มตัวอย่างมีขนาดเล็ก โดยที่ประชากรมีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ และความแปรปรวนเท่ากับ σ^2 จากบทที่ 5 หัวข้อที่ 5.4 จะได้ว่า

$$\bar{X}\sim N(\mu,\frac{\sigma^2}{n})$$

นั่นคือ

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$$

พิจารณาที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ และกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (โค้งปกติมาตรฐาน) ดังรูป



นั่นคือ

$$P(-Z_{\frac{\alpha}{2}}<Z<Z_{\frac{\alpha}{2}})=1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}}<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<Z_{\frac{\alpha}{2}})=1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\bar{X}-\mu<Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})=1-\alpha$$

พิจารณา

(1) กรณีที่ $\bar{X}-\mu > -Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ ดังนั้น

$$\bar{X}-\mu > -Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$-\mu > -\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$\mu < \bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

(2) กรณีที่ $\bar{X}-\mu < Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ ดังนั้น

$$\bar{X}-\mu < Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$-\mu < -\bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

$$\mu > \bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

จากทั้งสองกรณีสรุปได้ว่า

$$P(\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

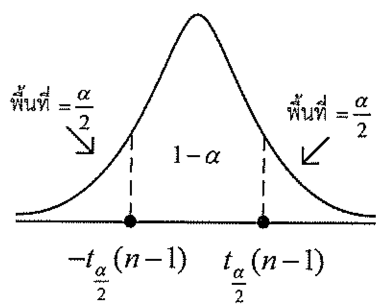
ดังนั้น ที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ ช่วงการประมาณค่าของ μ คือ

$$\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

พิจารณาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับค่าเฉลี่ยของกลุ่มตัวอย่าง เมื่อกลุ่มตัวอย่างมีขนาดเล็ก โดยที่ประชากรมีการแจกแจงแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ แต่ไม่ทราบความแปรปรวนของประชากรจากบทที่ 5 หัวข้อที่ 5.4 จะได้ว่า

$$T = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t(n-1)$$

พิจารณาที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ และกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบที ดังรูป



นั่นคือ

$$P(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-t_{\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{s/\sqrt{n}} < t_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \bar{X}-\mu < t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}) = 1-\alpha$$

พิจารณา

(1) กรณีที่ $\bar{X}-\mu > -t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$ ดังนั้น

$$\bar{X}-\mu > -t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$-\mu > -\bar{X}-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$\mu < \bar{X}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

(2) กรณีที่ $\bar{X}-\mu < t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$ ดังนั้น

$$\bar{X}-\mu < t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$-\mu < -\bar{X}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$\mu > \bar{X}-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

จากทั้งสองกรณีสรุปได้ว่า

$$P(\bar{X}-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \bar{X}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}) = 1-\alpha$$

ดังนั้น ที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ ช่วงการประมาณค่าของ μ คือ

$$\bar{X}-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \bar{X}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

การประมาณค่าแบบช่วงหรือการสร้างช่วงความเชื่อมั่นสำหรับค่าเฉลี่ยของประชากร กรณีกลุ่มตัวอย่างมีขนาดเล็ก ($n<30$) สามารถสรุปได้ดังนี้

กรณีที่ 2 การประมาณค่าเฉลี่ยของประชากรแบบช่วงกรณีกลุ่มตัวอย่างมีขนาดเล็ก ($n<30$)

เมื่อประชากรมีการแจกแจงความน่าจะเป็นแบบปกติด้วยค่าเฉลี่ยเท่ากับ μ

(1) ถ้าทราบค่าความแปรปรวนของประชากร σ^2 ดังนั้นช่วงการประมาณค่าของ μ ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

(2) ถ้าไม่ทราบค่าความแปรปรวนของประชากร σ^2 ดังนั้นช่วงการประมาณค่าของ μ ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$\bar{X}-t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}<\mu<\bar{X}+t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}}$$

เมื่อ μ หมายถึง ค่าเฉลี่ยของประชากร

$\bar{X}$ หมายถึง ค่าเฉลี่ยของกลุ่มตัวอย่าง

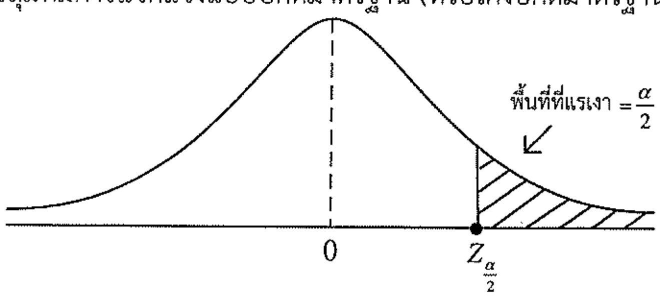
σ หมายถึง ส่วนเบี่ยงเบนมาตรฐานของประชากร

s หมายถึง ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

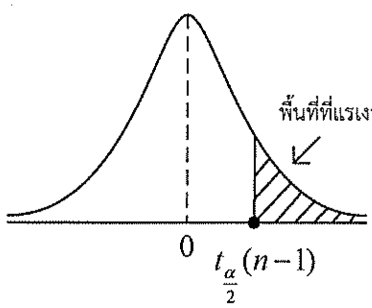
n หมายถึง ขนาดของกลุ่มตัวอย่าง

$1-\alpha$ หมายถึง ระดับความเชื่อมั่น

$Z_{\frac{\alpha}{2}}$ หมายถึง ค่าของตัวแปรสุ่ม $Z\sim N(0,1)$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน)



$t_{\frac{\alpha}{2}}(n-1)$ หมายถึง ค่าของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบที

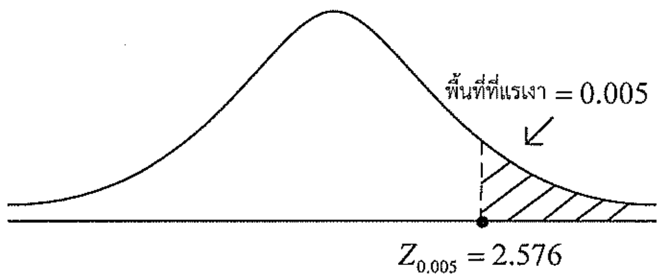


ตัวอย่างที่ 6.2.4 ผู้ผลิตช็อคโกแลตรายหนึ่งต้องการศึกษาเกี่ยวกับปริมาณแคลอรีที่ได้รับจากช็อคโกแลต 1 แท่ง จึงทำการสุ่มเลือกช็อคโกแลตมาทั้งหมด 9 แท่ง ปรากฏว่าปริมาณแคลอรีเฉลี่ยมีค่าเท่ากับ 220 แคลอรี ถ้าปริมาณแคลอรีที่ได้รับจากช็อคโกแลตมีการแจกแจงแบบปกติ จากประสบการณ์ของผู้ผลิตทราบว่าความแปรปรวนของปริมาณแคลอรีมีค่าเท่ากับ 40 จงหาช่วงความเชื่อมั่นของปริมาณแคลอรีเฉลี่ยที่ได้รับจากช็อคโกแลตที่ระดับความเชื่อมั่น 99%

วิธีทำ กำหนดให้ X แทนปริมาณแคลอรีที่ได้รับจากช็อคโกแลต 1 แท่ง จากโจทย์ ความแปรปรวนของประชากร ค่าเฉลี่ยของกลุ่มตัวอย่าง ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 40, 220, 9 และ 0.99 ตามลำดับ นั่นคือ $\sigma^2 = 40$, $\bar{X} = 220$, $n = 9$ และ $\alpha = 1 - 0.99 = 0.01$ ดังนั้น ปริมาณแคลอรีเฉลี่ยที่ได้รับจากช็อคโกแลตมีค่าอยู่ในช่วง

$$\bar{X} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ดังรูป



โดยใช้ตารางที่ 3.2 เมื่อ $F(z) = 1 - 0.005 = 0.995$ จะได้ว่า

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.01}{2}} = Z_{0.005} = 2.576$$

ดังนั้น

$$220 - 2.576 \frac{\sqrt{40}}{\sqrt{9}} < \mu < 220 + 2.576 \frac{\sqrt{40}}{\sqrt{9}}$$

$$214.5693 < \mu < 225.4307$$

ดังนั้น ปริมาณแคลอรีเฉลี่ยที่ได้รับจากช็อคโกแลตมีค่าประมาณ 214.5693 – 225.4307 แคลอรี ที่ระดับความเชื่อมั่น 99% ■

ตัวอย่างที่ 6.2.5 นักวิจัยคนหนึ่งต้องการศึกษาเกี่ยวกับปริมาณไขมันในไส้กรอก 1 ชิ้น จึงได้สุ่มตัวอย่างไส้กรอกมา 10 ชิ้น และนำไปตรวจหาปริมาณไขมันได้ดังนี้ (หน่วยเป็นมิลลิกรัม)

25.2, 21.3, 22.8, 17.0, 29.8, 21.0, 25.5, 16.0, 20.9, 19.5

โดยที่ปริมาณไขมันในไส้กรอกมีการแจกแจงแบบปกติ จงหาช่วงความเชื่อมั่นของปริมาณไขมันเฉลี่ยในไส้กรอกที่ระดับความเชื่อมั่น 95%

วิธีทำ กำหนดให้ X แทนปริมาณไขมันในไส้กรอก 1 ชิ้น จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 10 และ 0.95 ตามลำดับ นั่นคือ $n=10$ และ $\alpha=1-0.95=0.05$ ดังนั้น ปริมาณไขมันเฉลี่ยในไส้กรอกมีค่าอยู่ในช่วง

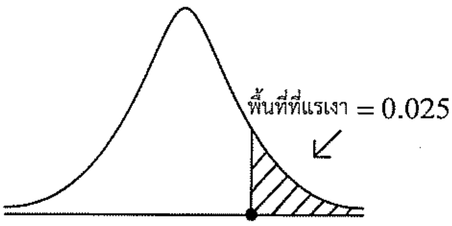
$$\bar{X}-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}<\mu<\bar{X}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

คำนวณหาค่า

$$\bar{X}=\frac{\sum_{i=1}^{n}X_i}{n}=\frac{25.2+\ldots+19.5}{10}=\frac{219}{10}=21.9$$

$$s^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}=\frac{(25.2-21.9)^2+\ldots+(19.5-21.9)^2}{9}=\frac{153.82}{9}\Rightarrow s=\sqrt{\frac{153.82}{9}}=\frac{\sqrt{153.82}}{3}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบที (หรือกราฟการแจกแจงแบบที) ซึ่งมีองศาอิสระเท่ากับ 9 ดังรูป



โดยใช้ตารางที่ 5 เมื่อ $\nu=9$ และ $p=1-0.025=0.975$ จะได้ว่า จะได้ว่า

$$t_{\frac{\alpha}{2}}(n-1)=t_{\frac{0.05}{2}}(10-1)=t_{0.025}(9)=2.262$$

ดังนั้น

$$21.9-2.262\frac{\frac{\sqrt{153.82}}{3}}{\sqrt{10}}<\mu<21.9+2.262\frac{\frac{\sqrt{153.82}}{3}}{\sqrt{10}}$$

$$18.9428<\mu<24.8572$$

ดังนั้น ปริมาณไขมันเฉลี่ยในไส้กรอกมีค่าประมาณ 18.9428 - 24.8572 มิลลิกรัม ที่ระดับความเชื่อมั่น 95% ∎

6.3 การประมาณค่าสัดส่วนของประชากร

6.3.1 การประมาณค่าสัดส่วนของประชากรแบบจุด

ประมาณค่าสัดส่วนประชากร (p) แบบจุดด้วยค่าสัดส่วนของกลุ่มตัวอย่าง ($\hat{p}$) โดยที่

$$p = \frac{N_p}{N}$$

$$\hat{p} = \frac{n_{\hat{p}}}{n}$$

เมื่อ N หมายถึง ขนาดประชากร

n หมายถึง ขนาดกลุ่มตัวอย่าง

N_p หมายถึง จำนวนของสิ่งที่สนใจทั้งหมดในประชากร

$n_{\hat{p}}$ หมายถึง จำนวนของสิ่งที่สนใจทั้งหมดในกลุ่มตัวอย่าง

p หมายถึง ค่าสัดส่วนของสิ่งที่สนใจในประชากร

$\hat{p}$ หมายถึง ค่าสัดส่วนของสิ่งที่สนใจในกลุ่มตัวอย่าง

พิจารณาการหาค่าเฉลี่ยของค่าสัดส่วนของกลุ่มตัวอย่าง ($\hat{p}$) ดังนี้

$$E(\hat{p}) = E\left(\frac{n_{\hat{p}}}{n}\right)$$

$$= \frac{1}{n}E\left(n_{\hat{p}}\right)$$

เนื่องจาก $n_{\hat{p}}$ หมายถึง จำนวนของสิ่งที่สนใจทั้งหมดในกลุ่มตัวอย่าง ดังนั้น $n_{\hat{p}}$ จึงเป็นตัวแปรสุ่มที่มีการแจกแจงแบบทวินาม ซึ่งค่าเฉลี่ยเท่ากับ np ทำให้ได้ว่า

$$E(\hat{p}) = \frac{np}{n} = p$$

เนื่องจาก $E(\hat{p}) = p$ จึงทำให้ $\hat{p}$ เป็นตัวประมาณค่าที่ไม่เอนเอียงของ p ดังนั้น สำหรับการประมาณค่าสัดส่วนของประชากร (p) แบบจุด จึงสามารถประมาณได้ด้วยค่าสัดส่วนของกลุ่มตัวอย่าง $(\hat{p})$ นั่นเอง

<u>**ตัวอย่างที่ 6.3.1**</u> มีคะแนนสอบของนักศึกษา 600 คน สุ่มคะแนนนักศึกษามา 10 คน พบว่ามีคะแนนดังนี้

36, 42, 34, 40, 64, 50, 22, 54, 60, 32

จงประมาณค่าสัดส่วนของนักศึกษา 600 คน ที่ได้คะแนนมากกว่า 50 คะแนน โดยใช้การประมาณค่าแบบจุด

วิธีทำ การประมาณค่าแบบจุดสำหรับค่าสัดส่วนของประชากร ทำได้โดยหาค่าสัดส่วนของกลุ่มตัวอย่างดังนี้

$$\hat{p} = \frac{n_{\hat{p}}}{n} = \frac{3}{10} = 0.3$$

เมื่อ $n_{\hat{p}}$ แทนจำนวนนักศึกษาจากกลุ่มตัวอย่างที่ได้คะแนนมากกว่า 50 คะแนน

ดังนั้น ค่าสัดส่วนของประชากรหรือค่าสัดส่วนของนักศึกษา 600 คน ที่ได้คะแนนมากกว่า 50 คะแนนมีค่าประมาณ 0.3 หรือ 30% ■

6.3.2 การประมาณค่าสัดส่วนของประชากรแบบช่วง

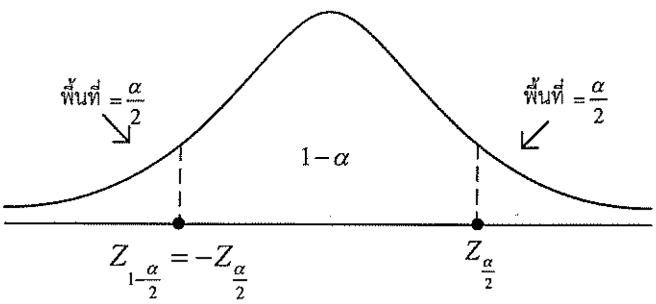
พิจารณาการแจกแจงความน่าจะเป็นสำหรับค่าสัดส่วนของกลุ่มตัวอย่าง เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ จากบทที่ 5 หัวข้อที่ 5.5 จะได้ว่า

$$\hat{p} \sim N(p, \frac{p(1-p)}{n})$$

นั่นคือ

$$Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \sim N(0,1)$$

พิจารณาที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ และกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (โค้งปกติมาตรฐาน) ดังรูป



นั่นคือ

$$P(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}} < \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} < Z_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(-Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} < \hat{p}-p < Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}) = 1-\alpha$$

พิจารณา

(1) กรณีที่ $\hat{p}-p > -Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$ ดังนั้น

$$\hat{p}-p > -Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

$$-p > -\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

$$p < \hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

(2) กรณีที่ $\hat{p}-p<Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$ ดังนั้น

$$\hat{p}-p<Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

$$-p<-\hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

$$p>\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}$$

จากทั้งสองกรณีสรุปได้ว่า

$$P(\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}<p<\hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}})=1-\alpha$$

เพื่อความสะดวกในการประมาณค่าจึงประมาณ p ด้วย $\hat{p}$ ดังนี้

$$P(\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}<p<\hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}})=1-\alpha$$

ดังนั้น ที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ ช่วงการประมาณค่าของ p คือ

$$\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}<p<\hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

การประมาณค่าแบบช่วงหรือการสร้างช่วงความเชื่อมั่นสำหรับค่าสัดส่วนของประชากร สามารถสรุปได้ดังนี้

การประมาณค่าสัดส่วนของประชากรแบบช่วง

ถ้ากลุ่มตัวอย่างมีขนาดใหญ่ ($n\geq 30$) ช่วงการประมาณค่าสำหรับค่าสัดส่วนของประชากร p ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}<p<\hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

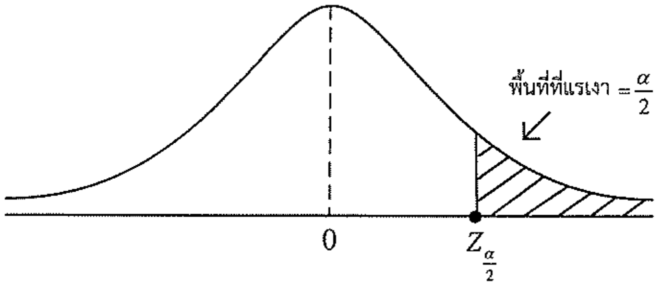
เมื่อ p หมายถึง ค่าสัดส่วนที่สนใจของประชากร

$\hat{p}$ หมายถึง ค่าสัดส่วนที่สนใจของกลุ่มตัวอย่าง

n หมายถึง ขนาดของกลุ่มตัวอย่าง

$1-\alpha$ หมายถึง ระดับความเชื่อมั่น

$Z_{\frac{\alpha}{2}}$ หมายถึง ค่าของตัวแปรสุ่ม $Z\sim N(0,1)$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน)



<u>**ตัวอย่างที่ 6.3.2**</u> การสำรวจเกี่ยวกับการมีบ้านของคนกรุงเทพฯ ได้มีการสุ่มคนกรุงเทพฯ มาจำนวน 100 คน พบว่าผู้ที่มีบ้านเป็นของตนเองมีจำนวน 60 คน

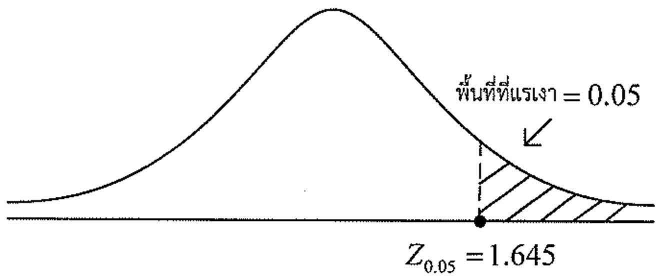
(1) จงประมาณค่าสัดส่วนของคนกรุงเทพฯ ที่มีบ้านเป็นของตนเอง ที่ระดับความเชื่อมั่น 90%

(2) จงประมาณค่าสัดส่วนของคนกรุงเทพฯ ที่ไม่มีบ้านเป็นของตนเอง ที่ระดับความเชื่อมั่น 95%

วิธีทำ (1) กำหนดให้ p แทนค่าสัดส่วนของคนกรุงเทพฯ ที่มีบ้านเป็นของตนเอง จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 100 และ 0.9 ตามลำดับ นั่นคือ $n=100$ และ $\alpha = 1-0.9 = 0.1$ เนื่องจากคนกรุงเทพฯ ที่มีบ้านเป็นของตนเองมีจำนวน 60 คน จากการสุ่มตัวอย่างมา 100 คน จะได้ว่า $\hat{p} = \frac{60}{100} = 0.6$ ดังนั้น ค่าสัดส่วนของคนกรุงเทพฯ ที่มีบ้านเป็นของตัวเองมีค่าอยู่ในช่วง

$$\hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ดังรูป



โดยใช้ตารางที่ 3.2 เมื่อ $F(z) = 1-0.05 = 0.95$ จะได้ว่า

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.1}{2}} = Z_{0.05} = 1.645$$

ดังนั้น

$$0.6 - 1.645\sqrt{\frac{(0.6)(1-0.6)}{100}} < p < 0.6 + 1.645\sqrt{\frac{(0.6)(1-0.6)}{100}}$$

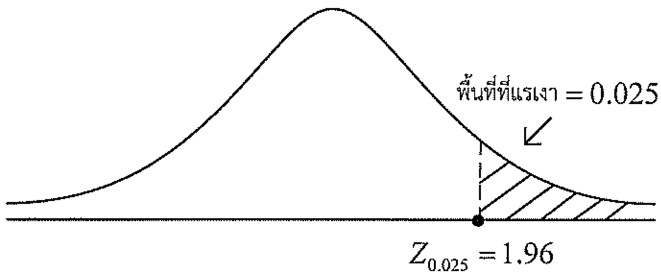
$$0.5194 < p < 0.6806$$

ดังนั้น ค่าสัดส่วนของคนกรุงเทพฯ ที่มีบ้านเป็นของตัวเองมีค่าประมาณ 0.5194 - 0.6806 ที่ระดับความเชื่อมั่น 90%

(2) กำหนดให้ q แทนค่าสัดส่วนของคนกรุงเทพฯ ที่ไม่มีบ้านเป็นของตนเอง จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 100 และ 0.95 ตามลำดับ นั่นคือ $n=100$ และ $\alpha=1-0.95=0.05$ เนื่องจากคนกรุงเทพฯ ที่มีบ้านเป็นของตนเองมีจำนวน 60 คน จากการสุ่มตัวอย่างมา 100 คน จะได้ว่า $\hat{q}=\dfrac{100-60}{100}=0.4$ ดังนั้น ค่าสัดส่วนของคนกรุงเทพฯ ที่ไม่มีบ้านเป็นของตัวเองมีค่าอยู่ในช่วง

$$\hat{q}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{q}(1-\hat{q})}{n}}<q<\hat{q}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{q}(1-\hat{q})}{n}}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ดังรูป



โดยใช้ตารางที่ 3.2 เมื่อ $F(z)=1-0.025=0.975$ จะได้ว่า

$$Z_{\frac{\alpha}{2}}=Z_{\frac{0.05}{2}}=Z_{0.025}=1.96$$

ดังนั้น

$$0.4-1.96\sqrt{\frac{(0.4)(1-0.4)}{100}}<q<0.4+1.96\sqrt{\frac{(0.4)(1-0.4)}{100}}$$

$$0.3040<p<0.4960$$

ดังนั้น ค่าสัดส่วนของคนกรุงเทพฯ ที่ไม่มีบ้านเป็นของตัวเองมีค่าประมาณ 0.3040 - 0.4960 ที่ระดับความเชื่อมั่น 95% ■

<u>ตัวอย่างที่ 6.3.3</u> การสำรวจเกี่ยวกับการติดตั้งจานดาวเทียมของครอบครัวไทย ได้มีการสุ่มครอบครัวมาทั้งหมด 400 ครอบครัว พบว่ามีครอบครัวที่ติดตั้งจานดาวเทียมจำนวน 260 ครอบครัว

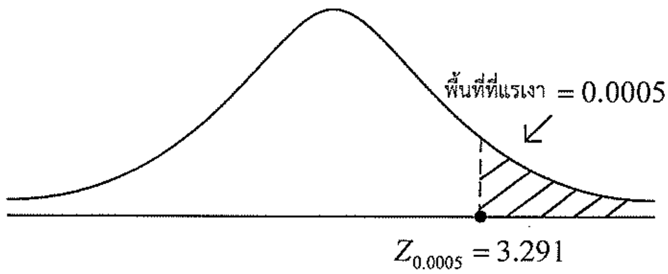
(1) จงประมาณค่าสัดส่วนของครอบครัวไทยที่ติดตั้งจานดาวเทียม ที่ระดับความเชื่อมั่น 99.9%

(2) จงประมาณค่าสัดส่วนของครอบครัวไทยที่ไม่ติดตั้งจานดาวเทียม ที่ระดับความเชื่อมั่น 98%

วิธีทำ (1) กำหนดให้ p แทนค่าสัดส่วนของครอบครัวไทยที่ติดตั้งจานดาวเทียม จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 400 และ 0.999 ตามลำดับ นั่นคือ $n=400$ และ $\alpha=1-0.999=0.001$ เนื่องจากครอบครัวไทยที่ติดตั้งจานดาวเทียมมีจำนวน 260 ครอบครัว จากการสุ่มตัวอย่างมา 400 ครอบครัว จะได้ว่า $\hat{p}=\dfrac{260}{400}=0.65$ ดังนั้น ค่าสัดส่วนของครอบครัวไทยที่ติดตั้งจานดาวเทียมมีค่าอยู่ในช่วง

$$\hat{p}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}<p<\hat{p}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ดังรูป



โดยใช้ตารางที่ 3.2 เมื่อ $F(z)=1-0.0005=0.9995$ จะได้ว่า

$$Z_{\frac{\alpha}{2}}=Z_{\frac{0.001}{2}}=Z_{0.0005}=3.291$$

ดังนั้น

$$0.65-3.291\sqrt{\frac{(0.65)(1-0.65)}{400}}<p<0.65+3.291\sqrt{\frac{(0.65)(1-0.65)}{400}}$$

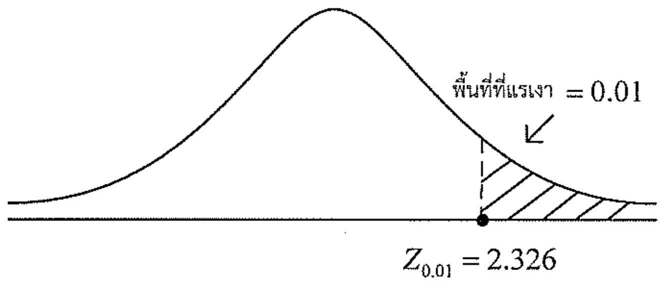
$$0.5715<p<0.7285$$

ดังนั้น ค่าสัดส่วนของครอบครัวไทยที่ติดตั้งจานดาวเทียมมีค่าประมาณ 0.5715 - 0.7285 ที่ระดับความเชื่อมั่น 99.9%

(2) กำหนดให้ q แทนค่าสัดส่วนของครอบครัวไทยที่ติดตั้งจานดาวเทียม จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 400 และ 0.98 ตามลำดับ นั่นคือ $n=400$ และ $\alpha=1-0.98=0.02$ เนื่องจากครอบครัวไทยที่ติดตั้งจานดาวเทียมมีจำนวน 260 ครอบครัว จากการสุ่มตัวอย่างมา 400 ครอบครัว จะได้ว่า $\hat{q}=\frac{400-260}{400}=0.35$ ดังนั้น ค่าสัดส่วนของครอบครัวไทยที่ไม่ติดตั้งจานดาวเทียมมีค่าอยู่ในช่วง

$$\hat{q}-Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{q}(1-\hat{q})}{n}}<q<\hat{q}+Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{q}(1-\hat{q})}{n}}$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ดังนี้



โดยใช้ตารางที่ 3.2 เมื่อ $F(z) = 1 - 0.01 = 0.99$ จะได้ว่า

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0.02}{2}} = Z_{0.01} = 2.326$$

ดังนั้น

$$0.35 - 2.326\sqrt{\frac{(0.35)(1-0.35)}{400}} < q < 0.35 + 2.326\sqrt{\frac{(0.35)(1-0.35)}{400}}$$

$$0.2945 < p < 0.4055$$

ดังนั้น ค่าสัดส่วนของครอบครัวไทยที่ไม่ติดตั้งจานดาวเทียมมีค่าประมาณ 0.2945 - 0.4055 ที่ระดับความเชื่อมั่น 98% ■

6.4 การประมาณค่าความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของประชากร

6.4.1 การประมาณค่าความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของประชากรแบบจุด

ประมาณค่าความแปรปรวนประชากร (σ^2) แบบจุดด้วยค่าความแปรปรวนของกลุ่มตัวอย่าง (s^2) โดยที่

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

เมื่อ X_i หมายถึง ค่าของสิ่งที่สนใจ
N หมายถึง ขนาดประชากร
n หมายถึง ขนาดกลุ่มตัวอย่าง
μ หมายถึง ค่าเฉลี่ยของประชากร
$\bar{X}$ หมายถึง ค่าเฉลี่ยของกลุ่มตัวอย่าง
σ^2 หมายถึง ความแปรปรวนของประชากร
s^2 หมายถึง ความแปรปรวนของกลุ่มตัวอย่าง

พิจารณาการหาค่าเฉลี่ยของความแปรปรวนของกลุ่มตัวอย่าง (s^2) ดังนี้

$$
\begin{aligned}
E(s^2) &= E\left(\sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n-1}\right)\\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}(X_i-\bar{X})^2\right)\\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}((X_i-\mu)-(\bar{X}-\mu))^2\right)\\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}\left((X_i-\mu)^2-2(X_i-\mu)(\bar{X}-\mu)+(\bar{X}-\mu)^2\right)\right)\\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}(X_i-\mu)^2-2(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\mu)+n(\bar{X}-\mu)^2\right)\\
&= \frac{1}{n-1}E\left(\sum_{i=1}^{n}(X_i-\mu)^2-n(\bar{X}-\mu)^2\right)\\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}E((X_i-\mu))^2-nE((\bar{X}-\mu))^2\right)\\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}V(X_i)-nV(\bar{X})\right)\\
&= \frac{1}{n-1}\left(\sum_{i=1}^{n}\sigma^2-n\left(\frac{\sigma^2}{n}\right)\right)\\
&= \frac{1}{n-1}(n\sigma^2-\sigma^2)=\sigma^2
\end{aligned}
$$

ดังนั้น

$$E(s^2)=\sigma^2$$

เนื่องจาก $E(s^2)=\sigma^2$ จึงทำให้ s^2 เป็นตัวประมาณค่าที่ไม่เอนเอียงของ σ^2 ดังนั้น สำหรับการประมาณค่าความแปรปรวนของประชากร (σ^2) แบบจุด จึงสามารถประมาณได้ด้วยความแปรปรวนของกลุ่มตัวอย่าง (s^2) นั่นเอง

<u>**ตัวอย่างที่ 6.4.1**</u> มีคะแนนสอบของนักศึกษา 600 คน สุ่มคะแนนนักศึกษามา 10 คน พบว่ามีคะแนนดังนี้

36, 42, 34, 40, 64, 50, 22, 54, 60, 32

จงประมาณค่าความแปรปรวนของคะแนนสอบนักศึกษา 600 คน โดยใช้การประมาณค่าแบบจุด

วิธีทำ การประมาณค่าแบบจุดสำหรับความแปรปรวนของประชากร ทำได้โดยหาความแปรปรวนของกลุ่มตัวอย่างดังนี้ โดยที่ $\bar{X}=43.4$ ได้มาจากตัวอย่างที่ 6.2.1

$$s^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$$

$$= \frac{\sum_{i=1}^{10}(X_i - 43.4)^2}{10-1}$$
$$= \frac{(36-43.4)^2 + (42-43.4)^2 + \ldots + (32-43.4)^2}{10}$$
$$= \frac{1600.4}{9}$$
$$\approx 177.82$$

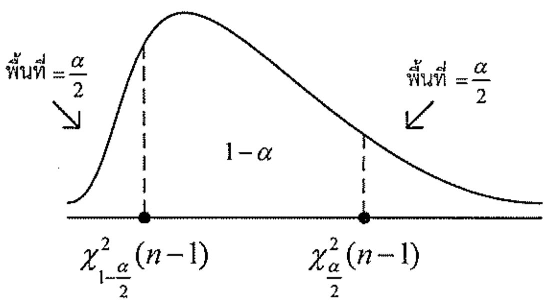
ดังนั้น ความแปรปรวนของประชากรหรือความแปรปรวนของคะแนนสอบนักศึกษา 600 คน มีค่าประมาณ 177.82 ■

6.4.2 การประมาณค่าความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของประชากรแบบช่วง

พิจารณาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่เกี่ยวข้องกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง เมื่อประชากรมีการแจกแจงแบบปกติ จากบทที่ 5 หัวข้อที่ 5.6 จะได้ว่า

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$$

พิจารณาที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ และกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ ดังรูป



นั่นคือ

$$P(-\chi^2_{1-\frac{\alpha}{2}}(n-1) < \chi^2 < \chi^2_{\frac{\alpha}{2}}(n-1)) = 1-\alpha$$
$$P(-\chi^2_{1-\frac{\alpha}{2}}(n-1) < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\frac{\alpha}{2}}(n-1)) = 1-\alpha$$
$$P(-\frac{\chi^2_{1-\frac{\alpha}{2}}(n-1)}{(n-1)s^2} < \frac{1}{\sigma^2} < \frac{\chi^2_{\frac{\alpha}{2}}(n-1)}{(n-1)s^2}) = 1-\alpha$$
$$P(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}) = 1-\alpha$$

ดังนั้น ที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ ช่วงการประมาณค่าของ σ^2 คือ

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$$

นั่นคือ ที่ระดับความเชื่อมั่นเท่ากับ $1-\alpha$ ช่วงการประมาณค่าของ σ^2 คือ

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}$$

การประมาณค่าแบบช่วงหรือการสร้างช่วงความเชื่อมั่นสำหรับความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของประชากร สามารถสรุปได้ดังนี้

การประมาณค่าความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของประชากรแบบช่วง

ถ้าประชากรมีการแจกแจงแบบปกติ ช่วงการประมาณค่าสำหรับความแปรปรวนของประชากร σ^2 ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$$

ถ้าประชากรมีการแจกแจงแบบปกติ ช่วงการประมาณค่าสำหรับส่วนเบี่ยงเบนมาตรฐานของประชากร σ ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ คือ

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}$$

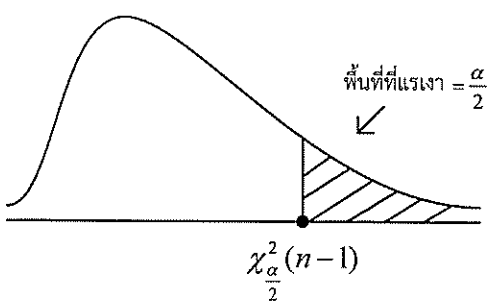
เมื่อ σ^2 หมายถึง ความแปรปรวนของประชากร

s^2 หมายถึง ความแปรปรวนของกลุ่มตัวอย่าง

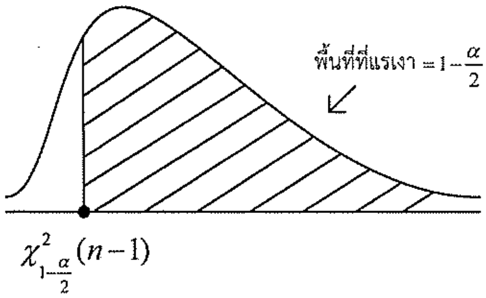
n หมายถึง ขนาดของกลุ่มตัวอย่าง

$1-\alpha$ หมายถึง ระดับความเชื่อมั่น

$\chi^2_{\frac{\alpha}{2}}(n-1)$ หมายถึง ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์



$\chi^2_{1-\frac{\alpha}{2}}(n-1)$ หมายถึง ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์



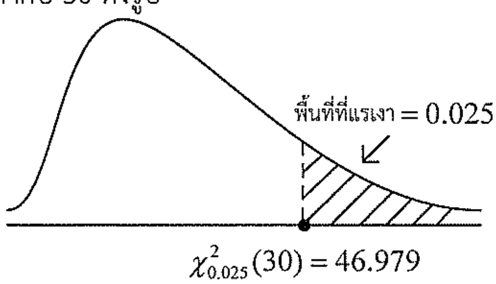
<u>ตัวอย่างที่ 6.4.2</u> ถ้าสุ่มตัวอย่างขนาด 31 จากประชากรที่มีการแจกแจงแบบปกติ และคำนวณความแปรปรวนของกลุ่มตัวอย่างได้เท่ากับ 25

(1) จงประมาณค่าความแปรปรวนของประชากร ที่ระดับความเชื่อมั่น 95%

(2) จงประมาณค่าส่วนเบี่ยงเบนมาตรฐานของประชากร ที่ระดับความเชื่อมั่น 99.9%

วิธีทำ (1) จากโจทย์ ขนาดของกลุ่มตัวอย่าง ความแปรปรวนของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 31, 25 และ 0.95 ตามลำดับ นั่นคือ $n=31$, $s^2=25$ และ $\alpha=1-0.95=0.05$

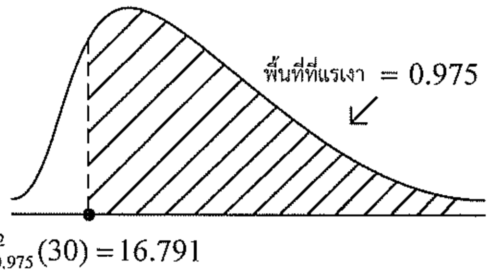
พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 30 ดังรูป



โดยใช้ตารางที่ 4.2 เมื่อ $v=30$ และ $p=1-0.025=0.975$ จะได้ว่า

$$\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{\frac{0.05}{2}}(31-1) = \chi^2_{0.025}(30) = 46.979$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 30 ดังรูป



โดยใช้ตารางที่ 4.1 เมื่อ $\nu = 30$ และ $p = 1 - 0.975 = 0.025$ จะได้ว่า

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{1-\frac{0.05}{2}}(31-1) = \chi^2_{0.975}(30) = 16.791$$

ความแปรปรวนของประชากรมีค่าอยู่ในช่วง

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$$

เมื่อแทนค่าตัวแปรทุกตัวลงไปจะได้ว่า

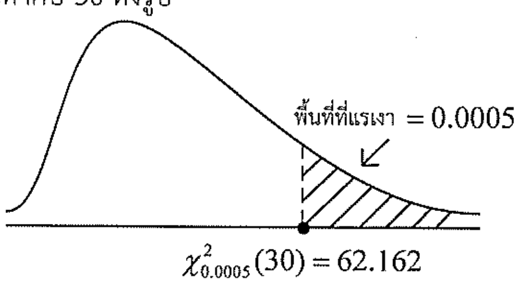
$$\frac{(31-1)25}{46.979} < \sigma^2 < \frac{(31-1)25}{16.791}$$

$$15.9646 < \sigma^2 < 44.6668$$

ดังนั้น ความแปรปรวนของประชากรมีค่าประมาณอยู่ในช่วง 15.9646 - 44.6668 ที่ระดับความเชื่อมั่น 95%

(2) จากโจทย์ ขนาดของกลุ่มตัวอย่าง ความแปรปรวนของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 31, 25 และ 0.999 ตามลำดับ นั่นคือ $n = 31$, $s^2 = 25$ และ $\alpha = 1 - 0.999 = 0.001$

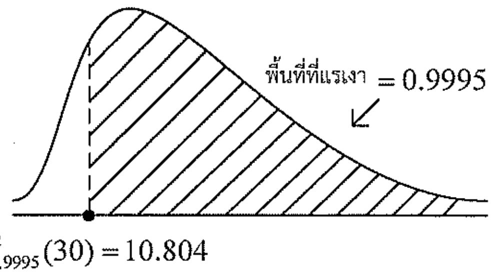
พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 30 ดังรูป



โดยใช้ตารางที่ 4.2 เมื่อ $\nu = 30$ และ $p = 1 - 0.0005 = 0.9995$ จะได้ว่า

$$\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{\frac{0.001}{2}}(31-1) = \chi^2_{0.0005}(30) = 62.162$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 30 ดังรูป



โดยใช้ตารางที่ 4.1 เมื่อ $\nu = 30$ และ $p = 1 - 0.9995 = 0.0005$ จะได้ว่า

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{1-\frac{0.001}{2}}(31-1) = \chi^2_{0.9995}(30) = 10.804$$

ส่วนเบี่ยงเบนมาตรฐานของประชากรมีค่าอยู่ในช่วง

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}$$

เมื่อแทนค่าตัวแปรทุกตัวลงไปจะได้ว่า

$$\sqrt{\frac{(31-1)25}{62.162}} < \sigma < \sqrt{\frac{(31-1)25}{10.804}}$$

$$3.4735 < \sigma < 8.3318$$

ดังนั้น ส่วนเบี่ยงเบนมาตรฐานของประชากรมีค่าประมาณอยู่ในช่วง 3.4735 – 8.3318 ที่ระดับความเชื่อมั่น 99.9% ■

<u>ตัวอย่างที่ 6.4.3</u> บริษัทไปรษณีย์ไทยทำการสำรวจน้ำหนักของพัสดุโดยการสุ่มพัสดุมาทั้งหมด 9 ชิ้น ซึ่งมีน้ำหนักดังนี้ (หน่วยเป็นกรัม)

30, 35, 58, 44, 21, 36, 19, 32, 48

(1) จงประมาณค่าความแปรปรวนของน้ำหนักพัสดุของบริษัทไปรษณีย์ไทย ที่ระดับความเชื่อมั่น 98%

(2) จงประมาณค่าส่วนเบี่ยงเบนมาตรฐานของน้ำหนักพัสดุของบริษัทไปรษณีย์ไทย ที่ระดับความเชื่อมั่น 99%

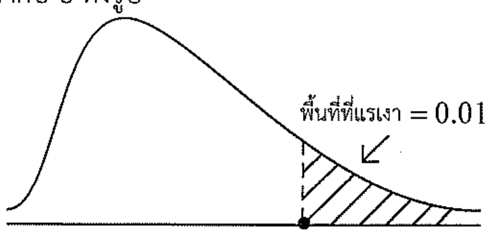
วิธีทำ (1) จากโจทย์ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 9 และ 0.98 ตามลำดับ นั่นคือ $n = 9$ และ $\alpha = 1 - 0.98 = 0.02$ ดังนั้น ความแปรปรวนของน้ำหนักพัสดุของบริษัทไปรษณีย์ไทยมีค่าอยู่ในช่วง

$$\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}$$

คำนวณหาค่า

$$s^2 = \frac{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}{n(n-1)} = \frac{9(12851) - (323)^2}{9(9-1)} = \frac{11330}{72} = \frac{5665}{36}$$

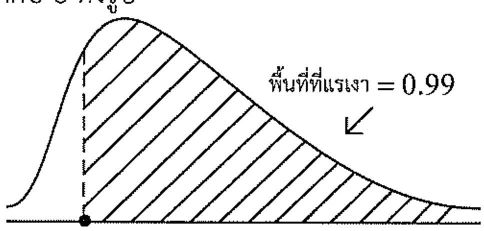
พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 8 ดังรูป



โดยใช้ตารางที่ 4.2 เมื่อ $v = 8$ และ $p = 1 - 0.01 = 0.99$ จะได้ว่า

$$\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{\frac{0.02}{2}}(9-1) = \chi^2_{0.01}(8) = 20.090$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 8 ดังรูป



โดยใช้ตารางที่ 4.1 เมื่อ $v = 8$ และ $p = 1 - 0.99 = 0.01$ จะได้ว่า

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{1-\frac{0.02}{2}}(31-1) = \chi^2_{0.99}(8) = 1.646$$

ดังนั้น

$$\frac{(9-1)\left(\frac{5665}{36}\right)}{20.090} < \sigma^2 < \frac{(9-1)\left(\frac{5665}{36}\right)}{1.646}$$

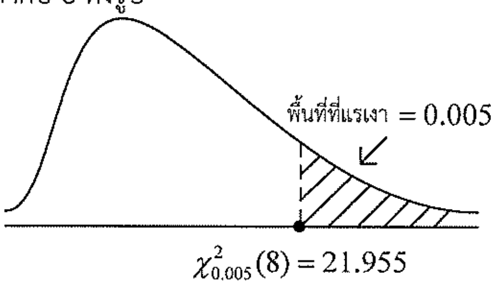
$$62.6625 < \sigma^2 < 764.8171$$

ดังนั้น ความแปรปรวนของน้ำหนักพัสดุของบริษัทไปรษณีย์ไทยมีค่าประมาณอยู่ในช่วง 62.6625 – 764.8171 ที่ระดับความเชื่อมั่น 98%

(2) จากโจทย์ขนาดของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 9 และ 0.99 ตามลำดับ นั่นคือ $n = 9$ และ $\alpha = 1 - 0.99 = 0.01$ ดังนั้น ส่วนเบี่ยงเบนมาตรฐานของน้ำหนักพัสดุของบริษัทไปรษณีย์ไทยมีค่าอยู่ในช่วง

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}}$$

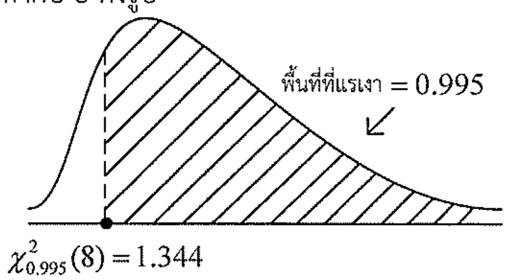
พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 8 ดังรูป



โดยใช้ตารางที่ 4.2 เมื่อ $\nu = 8$ และ $p = 1 - 0.005 = 0.995$ จะได้ว่า

$$\chi^2_{\frac{\alpha}{2}}(n-1) = \chi^2_{\frac{0.01}{2}}(9-1) = \chi^2_{0.005}(8) = 21.955$$

พิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ (หรือกราฟการแจกแจงแบบไคสแควร์) ซึ่งมีองศาอิสระเท่ากับ 8 ดังรูป



โดยใช้ตารางที่ 4.1 เมื่อ $\nu = 8$ และ $p = 1 - 0.995 = 0.005$ จะได้ว่า

$$\chi^2_{1-\frac{\alpha}{2}}(n-1) = \chi^2_{1-\frac{0.01}{2}}(31-1) = \chi^2_{0.995}(8) = 1.344$$

และจากข้อ (1) $s^2 = \dfrac{5665}{36}$ ดังนั้น

$$\sqrt{\frac{(9-1)\left(\frac{5665}{36}\right)}{21.955}} < \sigma < \sqrt{\frac{(9-1)\left(\frac{5665}{36}\right)}{1.344}}$$

$$7.5723 < \sigma < 30.6051$$

ดังนั้น ส่วนเบี่ยงเบนมาตรฐานของน้ำหนักพัสดุของบริษัทไปรษณีย์ไทยมีค่าประมาณอยู่ในช่วง 7.5723 – 30.6051 ที่ระดับความเชื่อมั่น 99% ■

บทที่ 7
การทดสอบสมมติฐานทางสถิติ
(Testing a Statistical Hypothesis)

ข้อมูลที่เกี่ยวข้องกับประชากรหรือพารามิเตอร์ สามารถประมาณค่าได้ด้วยการประมาณค่าแบบช่วงหรือการประมาณค่าแบบจุดตามที่กล่าวมาแล้วในบทที่ 6 แต่ในบางครั้งไม่จำเป็นต้องทราบค่าของพารามิเตอร์ แต่ต้องการทราบเพียงว่า สิ่งที่เราคาดเดาเกี่ยวกับพารามิเตอร์เป็นไปตามที่คาดการณ์เอาไว้หรือไม่ ดังนั้น ในบทที่ 7 จึงได้กล่าวถึงการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ย ค่าสัดส่วน และความแปรปรวน (หรือส่วนเบี่ยงเบนมาตรฐาน) ของประชากร ซึ่งจะช่วยให้สามารถตรวจสอบว่าค่าพารามิเตอร์เหล่านี้เป็นไปตามที่เชื่อหรือไม่

7.1 แนวคิดพื้นฐานของการทดสอบสมมติฐานทางสถิติ

7.1.1 ความหมายของสมมติฐานและการทดสอบสมมติฐานทางสถิติ

สมมติฐานทางสถิติ (statistic hypothesis) หมายถึง ความเชื่อหรือสิ่งที่คาดว่าจะเกิดขึ้นเกี่ยวกับพารามิเตอร์

การทดสอบสมมติฐาน (testing a statistical hypothesis) หมายถึง การทดสอบความเชื่อหรือสิ่งที่คาดไว้เกี่ยวกับพารามิเตอร์

7.1.2 การตั้งสมมติฐานทางสถิติ

การทดสอบสมมติฐานทางสถิติโดยใช้ข้อมูลที่ได้จากกลุ่มตัวอย่าง เพื่อสรุปว่าสิ่งที่คาดไว้เป็นจริงหรือไม่ จะต้องตั้งสมมติฐานเพื่อการทดสอบ ซึ่งประกอบไปด้วยสมมติฐาน 2 ชนิด ดังนี้

(1) สมมติฐานว่าง (Null Hypothesis) เขียนแทนได้ด้วยสัญลักษณ์ H_0

(2) สมมติฐานแย้ง (Alternative Hypothesis) เขียนแทนได้ด้วยสัญลักษณ์ H_1

7.1.3 หลักเกณฑ์ในการตั้งสมมติฐานทางสถิติ

การตั้งสมมติฐานทางสถิติจะต้องตั้งสมมติฐาน H_0 และ H_1 โดยมีหลักเกณฑ์ดังนี้

กรณีที่ 1: ถ้าสิ่งที่คาดไว้มีเครื่องหมายเท่ากับ ให้ตั้งเป็นสมมติฐาน H_0 ส่วนสมมติฐาน H_1 ให้ตั้งเป็นข้อความตรงข้ามกับสมมติฐาน H_0

กรณีที่ 2: ถ้าสิ่งที่คาดไว้ไม่มีเครื่องหมายเท่ากับ ให้ตั้งเป็นสมมติฐาน H_1 ส่วนสมมติฐาน H_0 ให้ตั้งเป็นข้อความตรงข้ามกับสมมติฐาน H_1

<u>ตัวอย่างที่ 7.1.1</u> ผู้จัดการโรงงานผลิตอาหารกระป๋องแห่งหนึ่งมีความเชื่อว่าน้ำหนักเฉลี่ยของอาหารกระป๋องที่ผลิตจะมีน้ำหนักอย่างน้อย 200 กรัม ถ้ากำหนดให้ μ แทนน้ำหนักเฉลี่ยของอาหารกระป๋องที่โรงงานแห่งนี้ผลิต นั่นคือ $\mu \geq 200$ ดังนั้น ถ้าต้องการทดสอบความเชื่อดังกล่าวจะสามารถตั้งสมมติฐานทางสถิติได้ดังนี้

$$H_0 : \mu \geq 200$$
$$H_1 : \mu < 200$$

<u>ตัวอย่างที่ 7.1.2</u> บริษัทผู้ผลิตเครื่องใช้ไฟฟ้าแห่งหนึ่งอ้างว่าเครื่องใช้ไฟฟ้าที่ผลิตมีอายุการใช้งานเฉลี่ยนานกว่า 1,000 ชั่วโมง ถ้ากำหนดให้ μ แทนอายุการใช้งานเฉลี่ยของเครื่องใช้ไฟฟ้าที่บริษัทแห่งนี้ผลิต นั่นคือ $\mu > 1,000$ ดังนั้น ถ้าต้องการทดสอบการกล่าวอ้างดังกล่าวจะสามารถตั้งสมมติฐานทางสถิติได้ดังนี้

$$H_0 : \mu \leq 1,000$$
$$H_1 : \mu > 1,000$$

<u>ตัวอย่างที่ 7.1.3</u> บริษัทผู้ผลิตหลอดไฟรายหนึ่งมีความเชื่อว่าหลอดไฟที่ผลิตจะชำรุดไม่เกิน 5% ถ้ากำหนดให้ p แทนค่าสัดส่วนของหลอดไฟชำรุดที่ผลิตจากบริษัทแห่งนี้ นั่นคือ $p \leq 0.05$ ดังนั้น ถ้าต้องการทดสอบความเชื่อดังกล่าวจะสามารถตั้งสมมติฐานทางสถิติได้ดังนี้

$$H_0 : p \leq 0.05$$
$$H_1 : p > 0.05$$

<u>ตัวอย่างที่ 7.1.4</u> บริษัทแห่งผลิตแบตเตอรีสำหรับรถยนต์แห่งหนึ่งมีความเชื่อว่าส่วนเบี่ยงเบนมาตรฐานของอายุการใช้งานแบตเตอรีที่ผลิตมีค่าเท่ากับ 0.9 ปี ถ้ากำหนดให้ σ แทนส่วนเบี่ยงเบนมาตรฐานของอายุการใช้งานแบตเตอรีที่ผลิตจากบริษัทแห่งนี้ นั่นคือ ดังนั้น $\sigma = 0.9$ ถ้าต้องการทดสอบความเชื่อดังกล่าวจะสามารถตั้งสมมติฐานทางสถิติได้ดังนี้

$$H_0 : \sigma = 0.9$$
$$H_1 : \sigma \neq 0.9$$

7.1.4 ความผิดพลาดที่เกิดขึ้นจากการทดสอบสมมติฐานทางสถิติ

การทดสอบสมมติฐานทางสถิติเป็นการนำข้อมูลที่ได้จากกลุ่มตัวอย่างเพื่อมาสรุปผลเกี่ยวกับประชากร ดังนั้น การทดสอบสมมติฐานจึงมีความผิดพลาดเกิดขึ้นเสมอไม่ว่าจะยอมรับ H_0 หรือปฏิเสธ H_0 ก็ตาม ความผิดพลาดที่เกิดขึ้นจากการทดสอบสมมติฐานทางสถิติแบ่งออกได้เป็น 2 แบบ ดังนี้

1. <u>ความผิดพลาดแบบที่ 1</u> (type I error) หมายถึง ความผิดพลาดที่เกิดจากการปฏิเสธ H_0 ทั้ง ๆ ที่ H_0 เป็นจริง ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 เรียกว่า ระดับนัยสำคัญ (level of significance) ซึ่งเขียนแทนได้ด้วยสัญลักษณ์ α ดังนั้น

$$\alpha = P\,(\text{ปฏิเสธ } H_0 \mid H_0 \text{เป็นจริง})$$

2. <u>ความผิดพลาดแบบที่ 2</u> (type II error) หมายถึง ความผิดพลาดที่เกิดจากการยอมรับ H_0 ทั้ง ๆ ที่ H_0 ไม่จริง เขียนแทนความน่าจะเป็นในการเกิดความผิดพลาดแบบที่ 2 ได้ด้วยสัญลักษณ์ β ดังนั้น

$$\beta = P\,(\text{ยอมรับ } H_0 \mid H_0 \text{เป็นไม่จริง})$$

การตัดสินใจ	**สถานการณ์ที่เกิดขึ้นจริง**	
	H_0 เป็นจริง	H_0 ไม่จริง
ปฏิเสธ H_0	ความผิดพลาดแบบที่ 1 (ความน่าจะเป็นเท่ากับ α)	การตัดสินใจถูกต้อง
ยอมรับ H_0	การตัดสินใจถูกต้อง	ความผิดพลาดแบบที่ 2 (ความน่าจะเป็นเท่ากับ β)

<u>**ตัวอย่างที่ 7.1.5**</u> บริษัทผลิตยารักษาโรคแห่งหนึ่งตั้งข้อกำหนดเกี่ยวกับมาตรฐานของยาที่ผลิตเอาไว้ว่า ยารักษาโรคสูตรใดที่ทำให้เกิดผลข้างเคียงในการรักษามากกว่า 50% จะต้องมีการปรับปรุงสูตรยาขึ้นมาใหม่ เพื่อตรวจสอบมาตรฐานยาชนิดหนึ่งจึงได้ทำการทดสอบกับคนไข้จำนวน 20 คน จงหา

(1) ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 (α) จากการทดสอบสมมติฐานทางสถิติ

(2) ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 2 หรือ (β) จากการทดสอบสมมติฐานทางสถิติ เมื่อ $p = 0.9$

วิธีทำ กำหนดให้ X แทนจำนวนคนไข้ที่ได้รับการรักษาด้วยยาชนิดนี้แล้วมีอาการข้างเคียงเกิดขึ้น จากโจทย์ขนาดของกลุ่มตัวอย่างมีค่าเท่ากับ 20 นั่นคือ $n = 20$ และสามารถตั้งสมมติฐานได้ดังนี้

$$H_0 : p \le 0.5$$
$$H_1 : p > 0.5$$

ถ้ากำหนดให้ $p = 0.5$ ดังนั้น X มีการแจกแจงความน่าจะเป็นแบบทวินามด้วย $n = 20$ และ $p = 0.5$ นั่นคือ $X \sim B(20, 0.5)$ โดยที่ $E(X) = np = 20 \times 0.5 = 10$ สำหรับคนไข้ที่ถูกทดสอบด้วยยาชนิดนี้ทุก ๆ 10 คน จาก 20 คน จะมีอาการข้างเคียงเกิดขึ้น นั่นคือ สมมติฐาน H_1 จะเป็นจริง เมื่อคนไข้ที่สุ่มมาทดสอบมีมากกว่า 10 คน ที่เกิดอาการข้างเคียง ในความเป็นจริงจะปฏิเสธ H_0 เมื่อ X มีจำนวนมากกว่า 10 ดังนั้น ในตัวอย่างนี้จึงกำหนดขอบเขตปฏิเสธสมมติฐานเป็น $\{X : 14 \le X \le 20\}$

การตัดสินใจ	สถานการณ์ที่เกิดขึ้นจริง	
	H_0 เป็นจริง ($p \le 0.5$) แต่ในที่นี้พิจารณา $p = 0.5$	H_0 ไม่จริง ($p > 0.5$) แต่ในที่นี้พิจารณา $p = 0.9$
ปฏิเสธ H_0 (เมื่อ $14 \le X \le 20$) ทำการปรับปรุงสูตรยาเพื่อลดผลข้างเคียงที่เกิดขึ้น	ความผิดพลาดแบบที่ 1 (ความน่าจะเป็นเท่ากับ α)	การตัดสินใจถูกต้อง
ยอมรับ H_0 (เมื่อ $0 \le X \le 13$) ไม่จำเป็นต้องปรับปรุงสูตรยา	การตัดสินใจถูกต้อง	ความผิดพลาดแบบที่ 2 (ความน่าจะเป็นเท่ากับ β)

(1) เนื่องจากความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 หรือ α เกิดขึ้นเมื่อปฏิเสธ H_0 ทั้ง ๆ ที่ H_0 เป็นจริง นั่นคือ จะปฏิเสธ H_0 เมื่อ $14 \le X \le 20$ และ H_0 เป็นจริงเมื่อ $p = 0.5$ หาได้ดังนี้

$$\begin{aligned}\alpha &= P\ (\text{ปฏิเสธ } H_0 \mid H_0 \text{เป็นจริง})\\ &= P(14 \le X \le 20 \mid p = 0.5)\\ &= P(X = 14 \mid p = 0.5) + \ldots + P(X = 20 \mid p = 0.5)\\ &= 0.0370 + 0.0148 + 0.0046 + 0.0011 + 0.0002 + 0 + 0\\ &= 0.0577\end{aligned}$$

ดังนั้น ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 หรือ α มีค่าเท่ากับ 0.0577

(2) ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 2 หรือ β เกิดขึ้นเมื่อยอมรับ H_0 ทั้ง ๆ ที่ H_0 ไม่จริง นั่นคือ จะปฏิเสธ H_0 เมื่อ $0 \le X \le 13$ และ H_0 เป็นจริงเมื่อ $p = 0.9$ หาได้ดังนี้

$$\begin{aligned}\beta &= P\ (\text{ยอมรับ } H_0 \mid H_0 \text{เป็นไม่จริง})\\ &= P(0 \le X \le 13 \mid p = 0.9)\\ &= P(X = 0 \mid p = 0.9) + \ldots + P(X = 13 \mid p = 0.9)\\ &= P(Y = 20 \mid p = 0.1) + \ldots + P(Y = 7 \mid p = 0.1)\\ &= 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0.0001 + 0.0004 + 0.0020\\ &= 0.0025\end{aligned}$$

ดังนั้น ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 2 หรือ β มีค่าเท่ากับ 0.0025 ■

<u>ตัวอย่างที่ 7.1.6</u> บริษัทผลิตชิ้นส่วนรถยนต์แห่งหนึ่งต้องการทดสอบกระบวนการผลิตแบบใหม่ โดยมีความเชื่อว่าชิ้นส่วนรถยนต์ที่ผลิตด้วยกระบวนการผลิตแบบใหม่จะสามารถทนต่อแรงอัดได้มากกว่ากระบวนการผลิตแบบเดิม นั่นคือ สามารถทนต่อแรงอัดได้เฉลี่ยมากกว่า 150 ปอนด์ต่อตารางนิ้ว ถ้าแรงอัดที่ชิ้นส่วนดังกล่าวสามารถทนได้มีการแจกแจงแบบปกติด้วยส่วนเบี่ยงเบนมาตรฐานเท่ากับ 25 ปอนด์ต่อตารางนิ้ว กำหนดให้ขนาดของกลุ่มตัวอย่างเท่ากับ 36 จงหา

(1) ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 (α) จากการทดสอบสมมติฐานทางสถิติ

(2) ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 2 (β) จากการทดสอบสมมติฐานทางสถิติ เมื่อ $\mu = 155$

วิธีทำ กำหนดให้ X แทนแรงอัดที่ชิ้นส่วนแต่ละชิ้นที่ผลิตสามารถทนได้ จากโจทย์ ส่วนเบี่ยงเบนมาตรฐานของประชากร และขนาดของกลุ่มตัวอย่างมีค่าเท่ากับ 25 และ 36 ตามลำดับ นั่นคือ $\sigma = 25$ และ $n = 36$ และสามารถตั้งสมมติฐานได้ดังนี้

$$H_0 : \mu \le 150$$
$$H_1 : \mu > 150$$

เนื่องจากค่าเฉลี่ยของกลุ่มตัวอย่าง $\bar{X}$ เป็นตัวประมาณค่าที่ไม่เอนเอียงของค่าเฉลี่ยประชากร μ ดังนั้น ถ้าต้องการพิจารณาว่าจะต้องเปลี่ยนกระบวนการผลิตเป็นแบบใหม่หรือไม่ จึงจะปฏิเสธ H_0 ถ้าค่าเฉลี่ยของกลุ่มตัวอย่าง $\bar{X}$ มีค่ามากเพียงพอที่จะปฏิเสธ H_0 ได้ เพราะการเปลี่ยนแปลงกระบวนการผลิตจะต้องใช้ค่าใช้จ่ายรวมถึงค่าแรงงานที่มากขึ้น ในตัวอย่างนี้จึงกำหนดขอบเขตปฏิเสธสมมติฐานเป็น $\{\bar{X} : \bar{X} > 152\}$

การตัดสินใจ	สถานการณ์ที่เกิดขึ้นจริง	
	H_0 เป็นจริง ($\mu \le 150$) แต่ในที่นี้พิจารณา $\mu = 150$	H_0 ไม่จริง ($\mu > 150$) แต่ในที่นี้พิจารณา $\mu = 155$
ปฏิเสธ H_0 (เมื่อ $\bar{X} > 152$) เปลี่ยนแปลงวิธีการผลิตเป็นแบบใหม่	ความผิดพลาดแบบที่ 1 (ความน่าจะเป็นเท่ากับ α)	การตัดสินใจถูกต้อง
ยอมรับ H_0 (เมื่อ $\bar{X} \le 152$) ใช้วิธีการผลิตแบบเดิม	การตัดสินใจถูกต้อง	ความผิดพลาดแบบที่ 2 (ความน่าจะเป็นเท่ากับ β)

(1) เนื่องจากความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 หรือ α เกิดขึ้นเมื่อปฏิเสธ H_0 ทั้ง ๆ ที่ H_0 เป็นจริง นั่นคือ จะปฏิเสธ H_0 เมื่อค่าเฉลี่ยของกลุ่มตัวอย่างตกอยู่ในขอบเขตปฏิเสธสมมติฐาน หรือ $\bar{X} > 152$ และ H_0 เป็นจริงเมื่อ $\mu = 150$ หาได้ดังนี้

$$\begin{aligned}
\alpha &= P\ (\text{ปฏิเสธ } H_0 \mid H_0 \text{เป็นจริง}) \\
&= P(\bar{X} > 152 \mid \mu = 150) \\
&= P(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{152 - \mu}{\sigma/\sqrt{n}} \mid \mu = 150) \\
&= P(\frac{\bar{X} - 150}{25/\sqrt{36}} > \frac{152 - 150}{25/\sqrt{36}}) \\
&= P(Z > 2.4) \\
&= 1 - P(Z \le 2.4) \\
&= 1 - 0.9981 \\
&= 0.0019
\end{aligned}$$

ดังนั้น ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 หรือ α มีค่าเท่ากับ 0.0019

(2) ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 2 หรือ β เกิดขึ้นเมื่อยอมรับ H_0 ทั้ง ๆ ที่ H_0 ไม่จริง นั่นคือ จะปฏิเสธ H_0 เมื่อค่าเฉลี่ยของกลุ่มตัวอย่างตกอยู่ในขอบเขตปฏิเสธสมมติฐาน หรือ $\bar{X} > 152$ และ H_0 เป็นจริงเมื่อ $\mu = 155$ หาได้ดังนี้

$$
\begin{aligned}
\beta &= P\,(\text{ยอมรับ } H_0 \mid H_0 \text{เป็นไม่จริง}) \\
&= P(\bar{X} \le 152 \mid \mu = 155) \\
&= P(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le \frac{152-\mu}{\sigma/\sqrt{n}} \mid \mu = 155) \\
&= P(\frac{\bar{X}-155}{25/\sqrt{36}} \le \frac{152-155}{25/\sqrt{36}}) \\
&= P(Z \le -0.72) \\
&= P(Z > 0.72) \\
&= 1 - P(Z \le 0.72) \\
&= 1 - 0.7642 \\
&= 0.2358
\end{aligned}
$$

ดังนั้น ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 2 หรือ β มีค่าเท่ากับ 0.2358 ■

<u>ข้อสังเกต</u> จากตัวอย่างที่ 7.1.5 และ 7.1.6

1. การคำนวณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 หรือ α ในตัวอย่างที่ 7.1.5 จะต้องสมมติให้ $p = 0.5$ ดังนั้น สมมติฐานอาจเขียนได้ดังนี้

$$
\begin{aligned}
H_0 &: p = 0.5 \\
H_1 &: p > 0.5
\end{aligned}
$$

เช่นเดียวกับการคำนวณความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 1 หรือ α ในตัวอย่างที่ 7.1.6 จะต้องสมมติให้ $\mu = 150$ ดังนั้น สมมติฐานอาจเขียนได้ดังนี้

$$
\begin{aligned}
H_0 &: \mu = 150 \\
H_1 &: \mu > 150
\end{aligned}
$$

2. ความน่าจะเป็นของการเกิดความผิดพลาดแบบที่ 2 หรือ β ในตัวอย่างที่ 7.1.5 และ 7.1.6 จะมีค่าขึ้นอยู่กับค่า p และ μ ตามลำดับ

7.1.5 ประเภทของการทดสอบสมมติฐานทางสถิติและขอบเขตปฏิเสธสมมติฐาน

พิจารณาการทดสอบสมมติฐานที่เกี่ยวข้องกับค่าพารามิเตอร์ θ ซึ่งประเภทของการทดสอบสมมติฐานจะมีความสัมพันธ์กับขอบเขตปฏิเสธสมมติฐาน โดยที่การทดสอบสมมติฐานทางสถิติแบ่งออกเป็น 2 แบบ ดังนี้

1. การทดสอบแบบข้างเดียว (one-sides test)}

ถ้าสมมติฐาน H_1 เป็นเครื่องหมายมากกว่าหรือน้อยกว่า จะเรียกการทดสอบนั้นว่าเป็นการทดสอบแบบข้างเดียว ดังนั้น รูปแบบของสมมติฐานประกอบไปด้วย

แบบที่ 1:

$$H_0 : \theta \le \theta_0$$
$$H_1 : \theta > \theta_0$$

ซึ่งจะมีขอบเขตปฏิเสธสมมติฐานคือ ค่าสถิติ $> U$ หรือมีขอบเขตปฏิเสธสมมติฐานอยู่ทางด้านขวา โดยมีขนาดของขอบเขตปฏิเสธสมมติฐานเท่ากับ ระดับนัยสำคัญ หรือ α

แบบที่ 2:

$$H_0 : \theta \ge \theta_0$$
$$H_1 : \theta < \theta_0$$

ซึ่งจะมีขอบเขตปฏิเสธสมมติฐานคือ ค่าสถิติ $< L$ หรือมีขอบเขตปฏิเสธสมมติฐานอยู่ทางด้านซ้าย โดยมีขนาดของขอบเขตปฏิเสธสมมติฐานเท่ากับ ระดับนัยสำคัญ หรือ α

2. การทดสอบแบบสองข้าง (two-sided test)

ถ้าสมมติฐาน H_1 เป็นเครื่องหมายไม่เท่ากับ จะเรียกการทดสอบนั้นว่าเป็นการทดสอบแบบสองข้าง ดังนั้น รูปแบบของสมมติฐานคือ

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \ne \theta_0$$

ซึ่งจะมีขอบเขตปฏิเสธสมมติฐานคือ ค่าสถิติ $< L$ หรือ ค่าสถิติ $> U$ หรือมีขอบเขตปฏิเสธสมมติฐานอยู่ทั้งทางด้านซ้ายและทางด้านขวา โดยมีขนาดของขอบเขตปฏิเสธสมมติฐานเท่ากับ ระดับนัยสำคัญ หรือ α

<u>หมายเหตุ</u>

บางครั้งการทดสอบสมมติฐานแบบข้างเดียวอาจเขียนได้ดังนี้

แบบที่ 1

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta > \theta_0$$

แบบที่ 2

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta < \theta_0$$

7.1.6 สรุปขั้นตอนของการทดสอบสมมติฐานทางสถิติ

ขั้นที่ 1: ตั้งสมมติฐาน

ขั้นที่ 2: เลือกค่าสถิติที่ใช้ในการทดสอบ

ขั้นที่ 3: คำนวณค่าสถิติที่ใช้ในการทดสอบ

ขั้นที่ 4: กำหนดระดับนัยสำคัญ

ขั้นที่ 5: สร้างขอบเขตปฏิเสธสมมติฐาน
ขั้นที่ 6: สรุปผลการทดสอบ

7.2 การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร

7.2.1 ค่าสถิติที่ใช้ในการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร

การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร

กรณีที่ 1: เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ หรือ $n \geq 30$

เมื่อประชากรมีการแจกแจงแบบใด ๆ

(1) ถ้าทราบค่าความแปรปรวนของประชากร หรือ σ^2 ดังนั้น ค่าสถิติที่ใช้ในการทดสอบสมมติฐานคือ

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

(2) ถ้าไม่ทราบค่าความแปรปรวนของประชากร หรือ σ^2 ดังนั้น ค่าสถิติที่ใช้ในการทดสอบสมมติฐานคือ

$$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

กรณีที่ 2: เมื่อกลุ่มตัวอย่างมีขนาดเล็ก หรือ $n < 30$

เมื่อประชากรมีการแจกแจงแบบปกติ

(1) ถ้าทราบค่าความแปรปรวนของประชากร หรือ σ^2 ดังนั้น ค่าสถิติที่ใช้ในการทดสอบสมมติฐานคือ

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

(2) ถ้าไม่ทราบค่าความแปรปรวนของประชากร หรือ σ^2 ดังนั้น ค่าสถิติที่ใช้ในการทดสอบสมมติฐานคือ

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

องศาอิสระ (degree of freedom) เท่ากับ $n-1$

เมื่อ μ หมายถึง ค่าเฉลี่ยของประชากร

$\bar{X}$ หมายถึง ค่าเฉลี่ยของกลุ่มตัวอย่าง

σ หมายถึง ส่วนเบี่ยงเบนมาตรฐานของประชากร

s หมายถึง ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

n หมายถึง ขนาดของกลุ่มตัวอย่าง

7.2.2 ขอบเขตปฏิเสธสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร

จากค่าสถิติที่ใช้ทดสอบสามารถสรุปขอบเขตปฏิเสธสมมติฐานเกี่ยวกับค่าเฉลี่ยประชากรได้ดังนี้

การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร

1. การทดสอบแบบข้างเดียว

แบบที่ 1:

$$H_0 : \mu \le \mu_0$$
$$H_1 : \mu > \mu_0$$

ขอบเขตปฏิเสธสมมติฐานคือ $Z > Z_\alpha$ หรือ $t > t_\alpha(n-1)$ ขึ้นอยู่กับค่าสถิติที่ใช้ในการทดสอบสมมติฐาน

แบบที่ 2:

$$H_0 : \mu \ge \mu_0$$
$$H_1 : \mu < \mu_0$$

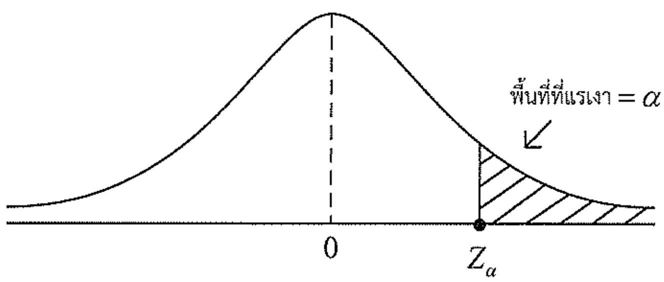
ขอบเขตปฏิเสธสมมติฐานคือ $Z < -Z_\alpha$ หรือ $t < -t_\alpha(n-1)$ ขึ้นอยู่กับค่าสถิติที่ใช้ในการทดสอบสมมติฐาน

2. การทดสอบแบบสองข้าง

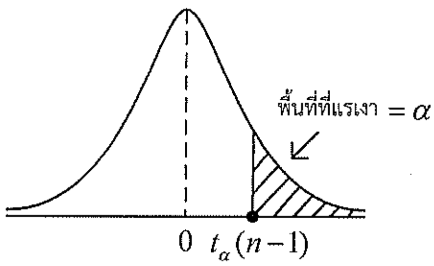
$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \ne \mu_0$$

ขอบเขตปฏิเสธสมมติฐานคือ $|Z| > Z_{\frac{\alpha}{2}}$ หรือ $|\,t\,| > t_{\frac{\alpha}{2}}(n-1)$ ขึ้นอยู่กับค่าสถิติที่ใช้ในการทดสอบสมมติฐาน

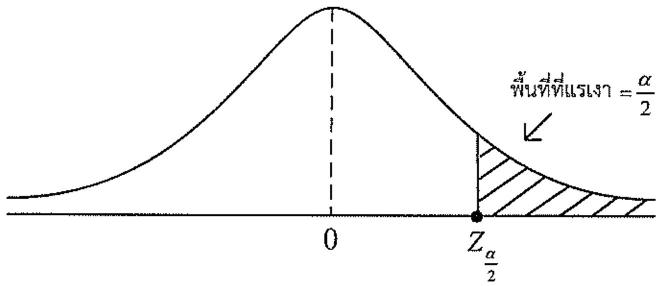
เมื่อ Z_α หมายถึง ค่าของตัวแปรสุ่ม $Z \sim N(0,1)$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน)



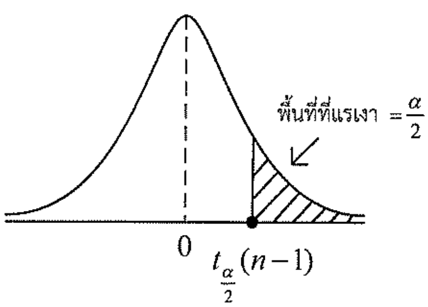
$t_\alpha(n-1)$ หมายถึง ค่าของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบที



$Z_{\frac{\alpha}{2}}$ หมายถึง ค่าของตัวแปรสุ่ม $Z \sim N(0,1)$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน)



$t_{\frac{\alpha}{2}}(n-1)$ หมายถึง ค่าของตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบที



<u>ตัวอย่างที่ 7.2.1</u> เจ้าของหอพักแห่งหนึ่งมีความเชื่อว่า ผู้เช่าห้องจะใช้ไฟฟ้าเฉลี่ยเดือนละไม่เกิน 45 หน่วยต่อห้อง เขาจึงได้สุ่มผู้เช่าห้องมาจำนวน 36 ห้อง เพื่อสอบถามจำนวนหน่วยไฟฟ้าที่ใช้ในเดือนนี้พบว่ามีค่าเฉลี่ยเท่ากับ 47 หน่วย ถ้าอัตราการใช้ไฟฟ้าแต่ละเดือนของผู้เช่าห้องในหอพักแห่งนี้มีการแจกแจงแบบปกติด้วยส่วนเบี่ยงเบนมาตรฐานเท่ากับ 9 ความเชื่อของเจ้าของหอพักนี้สมเหตุสมผลหรือไม่ ที่ระดับนัยสำคัญ 0.001

วิธีทำ กำหนดให้ X แทนอัตราการใช้ไฟฟ้าแต่ละเดือนของผู้เช่าห้องในหอพักแห่งนี้ จากโจทย์ ค่าเฉลี่ยของกลุ่มตัวอย่าง ส่วนเบี่ยงเบนมาตรฐานของประชากร ขนาดของกลุ่มตัวอย่าง และระดับนัยสำคัญมีค่าเท่ากับ 44, 9, 36 และ 0.001 ตามลำดับ นั่นคือ $\overline{X}=44$, $\sigma=9$, $n=36$ และ $\alpha=0.001$

<u>ขั้นที่ 1:</u> ตั้งสมมติฐาน

เนื่องจากเจ้าของหอพักแห่งนี้มีความเชื่อว่า ผู้เช่าห้องจะใช้ไฟฟ้าเฉลี่ยเดือนละไม่เกิน 45 หน่วยต่อห้อง นั่นคือ $\mu \le 45$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0 : \mu \le 45$$
$$H_1 : \mu > 45$$

ขั้นที่ 2: เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจาก $n \ge 30$ และทราบค่าความแปรปรวนของประชากร ดังนั้น จึงเลือกค่าสถิติ

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

ขั้นที่ 3: คำนวณค่าสถิติ

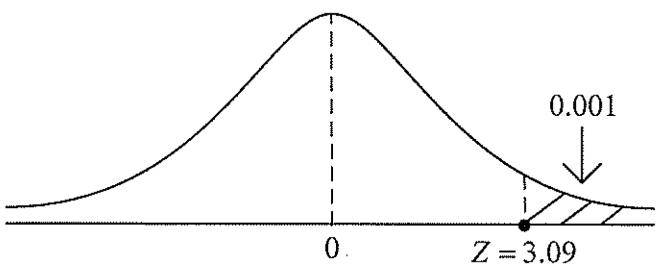
$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{47-45}{9/\sqrt{36}} = 1.33$$

ขั้นที่ 4: กำหนดระดับนัยสำคัญ

$$\alpha = 0.001$$

ขั้นที่ 5: สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : \mu > 45$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านขวาของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha = 0.001$ ดังรูป



โดยใช้ตารางที่ 3.2 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $Z > 3.09$

ขั้นที่ 6: สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($Z = 1.33$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($Z > 3.09$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.001

จากข้อมูลที่ให้มาข้างต้น สามารถกล่าวได้ว่าผู้เช่าห้องในหอพักแห่งนี้ใช้ไฟฟ้าเฉลี่ยเดือนละไม่เกิน 45 หน่วยต่อห้องที่ระดับนัยสำคัญ 0.001 นั่นคือความเชื่อของเจ้าของหอพักแห่งนี้สมเหตุสมผล ที่ระดับนัยสำคัญ 0.001 ■

<u>ตัวอย่างที่ 7.2.2</u> ผู้จัดการโรงงานอุตสาหกรรมแห่งหนึ่ง คาดว่าปริมาณวัตถุดิบเฉลี่ยที่ใช้ในโรงงานจะไม่ต่ำกว่า 880 ตันต่อวัน เพื่อที่จะทดสอบความเชื่อดังกล่าว จึงเก็บข้อมูลปริมาณวัตถุดิบที่ใช้ต่อวันมา 50 วัน คำนวณปริมาณวัตถุดิบเฉลี่ยได้ 871 ตันต่อวัน ส่วนเบี่ยงเบนมาตรฐานมีค่าเท่ากับ 21 ตัน การคาดคะเนของผู้จัดการถูกต้องหรือไม่ ที่ระดับนัยสำคัญ 0.05

วิธีทำ กำหนดให้ X แทนปริมาณวัตถุดิบที่โรงงานแห่งนี้ใช้ในแต่ละวัน จากโจทย์ ค่าเฉลี่ยของกลุ่มตัวอย่าง ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง ขนาดของกลุ่มตัวอย่าง และระดับนัยสำคัญมีค่าเท่ากับ 871, 21, 50 และ 0.05 ตามลำดับ นั่นคือ $\bar{X}=871,\ s=21,\ n=50$ และ $\alpha=0.05$

<u>ขั้นที่ 1:</u> ตั้งสมมติฐาน

เนื่องจากผู้จัดการโรงงานอุตสาหกรรมแห่งหนึ่ง คาดว่าปริมาณวัตถุดิบเฉลี่ยที่ใช้ในโรงงานจะไม่ต่ำกว่า 880 ตันต่อวัน นั่นคือ $\mu \ge 880$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0 : \mu \ge 880$$
$$H_1 : \mu < 880$$

<u>ขั้นที่ 2:</u> เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจาก $n \ge 30$ และไม่ทราบค่าความแปรปรวนของประชากร ดังนั้น จึงเลือกค่าสถิติ

$$Z=\frac{\bar{X}-\mu}{s/\sqrt{n}}$$

<u>ขั้นที่ 3:</u> คำนวณค่าสถิติ

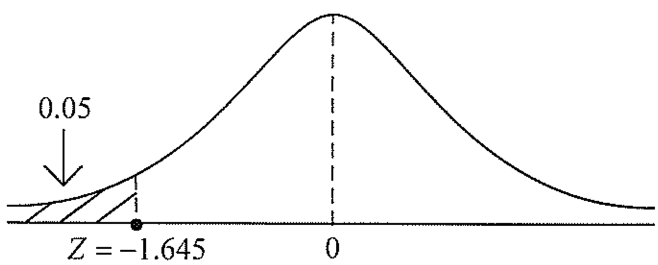
$$Z=\frac{\bar{X}-\mu}{s/\sqrt{n}}=\frac{871-880}{21/\sqrt{50}}=-3.03$$

<u>ขั้นที่ 4:</u> กำหนดระดับนัยสำคัญ

$$\alpha=0.05$$

<u>ขั้นที่ 5:</u> สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : \mu < 880$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านซ้ายของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha=0.05$ ดังรูป



โดยใช้ตารางที่ 3.2 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $Z < -1.645$

ขั้นที่ 6: สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($Z = -3.03$) ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($Z < -1.645$) ดังนั้นจึงปฏิเสธ H_0 ที่ระดับนัยสำคัญ 0.05

จากข้อมูลที่ให้มาข้างต้น ไม่สามารถกล่าวได้ว่าปริมาณวัตถุดิบเฉลี่ยที่ใช้ในโรงงานมีไม่ต่ำกว่า 880 ตันต่อวัน ที่ระดับนัยสำคัญ 0.05 นั่นคือการคาดคะเนของผู้จัดการไม่ถูกต้องที่ระดับนัยสำคัญ 0.05 ■

ตัวอย่างที่ 7.2.3 โรงงานผลิตอาหารกระป๋องแห่งหนึ่งกล่าวอ้างว่าน้ำหนักเฉลี่ยของอาหารกระป๋องที่ผลิตมีค่าไม่ต่ำกว่า 10 กรัม เพื่อทำการตรวจสอบการกล่าวอ้างของโรงงานแห่งนี้จึงได้ทำการสุ่มตัวอย่างอาหารกระป๋องมา 9 กระป๋อง ซึ่งมีน้ำหนักดังต่อไปนี้

10.1 9.9 10.2 9.6 9.5 10.2 9.4 9.8 10.4

ถ้าน้ำหนักอาหารกระป๋องที่ผลิตจากโรงงานแห่งนี้มีการแจกแจงแบบปกติด้วยส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.2 กรัม การกล่าวอ้างของโรงงานแห่งนี้น่าเชื่อถือหรือไม่ ที่ระดับนัยสำคัญ 0.5%

วิธีทำ กำหนดให้ X แทนน้ำหนักอาหารกระป๋องที่ผลิตจากโรงงานแห่งนี้ จากโจทย์ ส่วนเบี่ยงเบนมาตรฐานของประชากร ขนาดของกลุ่มตัวอย่าง และระดับนัยสำคัญมีค่าเท่ากับ 0.2, 9 และ 0.005 ตามลำดับ นั่นคือ $\sigma = 0.2$, $n = 9$ และ $\alpha = 0.005$ คำนวณหาค่าเฉลี่ยของกลุ่มตัวอย่างได้ดังนี้

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{10.1 + \ldots + 10.4}{9} = \frac{89.1}{9} = 9.9$$

ขั้นที่ 1: ตั้งสมมติฐาน

เนื่องจากโรงงานผลิตอาหารกระป๋องแห่งหนึ่งกล่าวอ้างว่าน้ำหนักเฉลี่ยของอาหารกระป๋องที่ผลิตมีค่าไม่ต่ำกว่า 10 กรัม นั่นคือ $\mu \geq 10$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0 : \mu \geq 10$$
$$H_1 : \mu < 10$$

ขั้นที่ 2: เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจาก $n < 30$ และทราบค่าความแปรปรวนของประชากร ดังนั้น จึงเลือกค่าสถิติ

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

ขั้นที่ 3: คำนวณค่าสถิติ

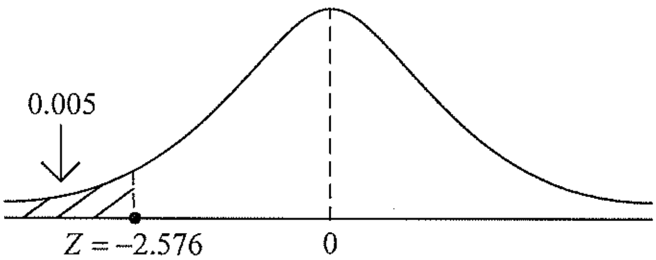
$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{9.9 - 10}{0.2 / \sqrt{9}} = -1.5$$

ขั้นที่ 4: กำหนดระดับนัยสำคัญ

$$\alpha = 0.005$$

ขั้นที่ 5: สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : \mu < 10$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านซ้ายของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha = 0.005$ ดังรูป



โดยใช้ตารางที่ 3.2 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $Z < -2.576$

ขั้นที่ 6: สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($Z = -1.5$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($Z < -2.576$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.005

จากข้อมูลที่ให้มาข้างต้น สามารถกล่าวได้ว่าน้ำหนักเฉลี่ยของอาหารกระป๋องที่ผลิตมีค่าไม่ต่ำกว่า 10 กรัม ที่ระดับนัยสำคัญ 0.001 นั่นคือการกล่าวอ้างของโรงงานแห่งนี้น่าเชื่อถือ ที่ระดับนัยสำคัญ 0.005

■

ตัวอย่างที่ 7.2.4 ผู้ผลิตไอศครีมรายหนึ่งเชื่อว่าไอศครีมที่ผลิตมีแคลอรี่เฉลี่ย 500 แคลอรี่ต่อกรัม เพื่อที่จะทดสอบความเชื่อนี้ เขาจึงสุ่มไอศครีมหนักก้อนละ 1 กรัม มา 25 ก้อน ทำการวัดปริมาณแคลอรี่ของไอศครีมแต่ละก้อนได้ดังนี้

505	520	565	550	490	480	480	540	550	440
520	550	580	500	520	510	460	490	490	480
520	530	500	480	500					

อยากทราบว่าความเชื่อของผู้ผลิตสมเหตุสมผลหรือไม่ ที่ระดับนัยสำคัญ 0.10 กำหนดให้ปริมาณแคลอรี่ในไอศครีมหนัก 1 กรัม มีการแจกแจงแบบปกติ

วิธีทำ กำหนดให้ X แทนปริมาณแคลอรีในไอศกรีมหนัก 1 กรัม จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับนัยสำคัญมีค่าเท่ากับ 25 และ 0.10 ตามลำดับ นั่นคือ $n=25$ และ $\alpha=0.1$ คำนวณหาค่าเฉลี่ยของกลุ่มตัวอย่าง และส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างได้ดังนี้

$$\bar{X}=\frac{\sum_{i=1}^{n}X_i}{n}=\frac{505+\ldots+500}{25}=\frac{12750}{25}=510$$

$$s^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}=\frac{(505-510)^2+\ldots+(500-510)^2}{25-1}=\frac{26950}{24}=\frac{13475}{12}\Rightarrow s=\sqrt{\frac{13475}{12}}$$

<u>ขั้นที่ 1:</u> ตั้งสมมติฐาน

เนื่องจากผู้ผลิตไอศครีมรายหนึ่งเชื่อว่าไอศครีมที่ผลิตมีแคลอรี่เฉลี่ย 500 แคลอรี่ต่อกรัม นั่นคือ $\mu=500$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0:\mu=500$$
$$H_1:\mu\neq 500$$

<u>ขั้นที่ 2:</u> เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจาก $n<30$ และไม่ทราบค่าความแปรปรวนของประชากร ดังนั้น จึงเลือกค่าสถิติ

$$t=\frac{\bar{X}-\mu}{s/\sqrt{n}}$$

โดยที่ องศาอิสระเท่ากับ $n-1$

<u>ขั้นที่ 3:</u> คำนวณค่าสถิติ

$$t=\frac{\bar{X}-\mu}{s/\sqrt{n}}=\frac{510-500}{\sqrt{\frac{13475}{12}}\Big/\sqrt{25}}=1.492$$

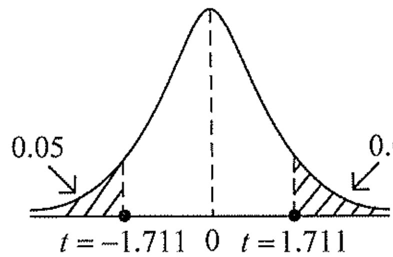
โดยที่ องศาอิสระเท่ากับ $n-1=25-1=24$

<u>ขั้นที่ 4:</u> กำหนดระดับนัยสำคัญ

$$\alpha=0.1$$

<u>ขั้นที่ 5:</u> สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบที จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบที ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1:\mu\neq 500$ จึงทำให้เป็นการทดสอบสมมติฐานแบบสองข้าง โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านซ้ายและขวาของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha=0.1$ ดังรูป



โดยใช้ตารางที่ 5 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบที จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $t<-1.711$ หรือ $t>1.711$

ขั้นที่ 6: สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($t=1.492$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($t<-1.711$ หรือ $t>1.711$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.1

จากข้อมูลที่ให้มาข้างต้น สามารถกล่าวได้ว่าแคลอรีเฉลี่ยในไอศครีมที่ผลิตมีค่าเท่ากับ 500 แคลอรีต่อกรัม ที่ระดับนัยสำคัญ 0.1 นั่นคือ ความเชื่อของผู้ผลิตไอศครีมรายนี้สมเหตุสมผลที่ระดับนัยสำคัญ 0.1 ■

7.3 การทดสอบสมมติฐานเกี่ยวกับค่าสัดส่วนของประชากร

7.3.1 ค่าสถิติที่ใช้ในการทดสอบสมมติฐานเกี่ยวกับค่าสัดส่วนของประชากร

การทดสอบสมมติฐานเกี่ยวกับค่าสัดส่วนของประชากร

เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ หรือ $n \geq 30$ ดังนั้น ค่าสถิติที่ใช้ในการทดสอบสมมติฐานคือ

$$Z=\frac{\hat{p}-p}{\sqrt{p(1-p)/n}}$$

เมื่อ p หมายถึง ค่าสัดส่วนที่สนใจของประชากร

$\hat{p}$ หมายถึง ค่าสัดส่วนที่สนใจของกลุ่มตัวอย่าง

n หมายถึง ขนาดของกลุ่มตัวอย่าง

7.3.2 ขอบเขตปฏิเสธสมมติฐานเกี่ยวกับค่าสัดส่วนของประชากร

จากค่าสถิติที่ใช้ทดสอบสามารถสรุปขอบเขตปฏิเสธสมมติฐานเกี่ยวกับค่าสัดส่วนของประชากรได้ดังนี้

การทดสอบสมมติฐานเกี่ยวกับค่าสัดส่วนของประชากร

1. การทดสอบแบบข้างเดียว

แบบที่ 1:

$$H_0 : p \le p_0$$
$$H_1 : p > p_0$$

ขอบเขตปฏิเสธสมมติฐานคือ $Z > Z_\alpha$

แบบที่ 2:

$$H_0 : p \ge p_0$$
$$H_1 : p < p_0$$

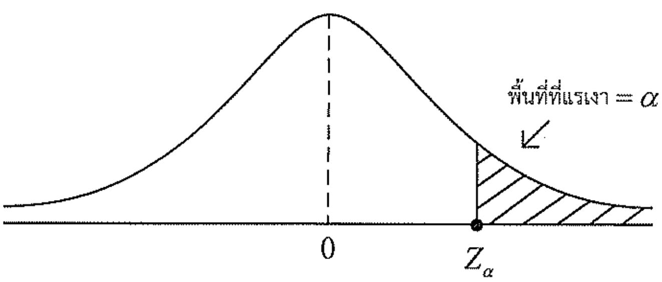
ขอบเขตปฏิเสธสมมติฐานคือ $Z < -Z_\alpha$

2. การทดสอบแบบสองข้าง

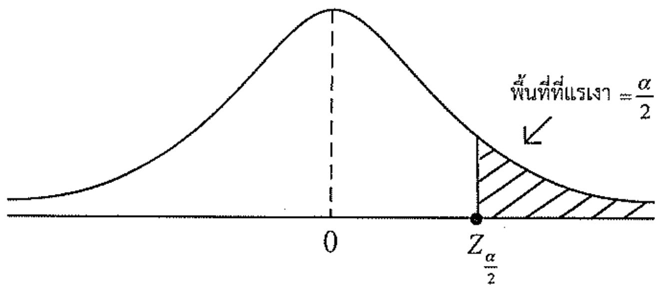
$$H_0 : p = p_0$$
$$H_1 : p \ne p_0$$

ขอบเขตปฏิเสธสมมติฐานคือ $|Z| > Z_{\frac{\alpha}{2}}$

เมื่อ Z_α หมายถึง ค่าของตัวแปรสุ่ม $Z \sim N(0,1)$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน)



$Z_{\frac{\alpha}{2}}$ หมายถึง ค่าของตัวแปรสุ่ม $Z \sim N(0,1)$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน)



<u>ตัวอย่างที่ 7.3.1</u> บริษัทขายเครื่องสำอางยี่ห้อหนึ่งมีความเชื่อว่าผู้หญิงไทยใช้เครื่องสำอางยี่ห้อนี้อย่างน้อย 20% เพื่อที่จะทดสอบความเชื่อนี้ จึงสุ่มตัวอย่างผู้หญิงไทยมา 500 คน ปรากฏว่ามีผู้ใช้เครื่องสำอางยี่ห้อนี้จำนวน 95 คน อยากทราบว่าความเชื่อของบริษัทนี้สมเหตุสมผลหรือไม่ ที่ระดับนัยสำคัญ 0.01

วิธีทำ กำหนดให้ p แทนค่าสัดส่วนของผู้หญิงไทยที่ใช้เครื่องสำอางยี่ห้อนี้ จากโจทย์ ขนาดของกลุ่มตัวอย่างและระดับนัยสำคัญมีค่าเท่ากับ 500 และ 0.01 ตามลำดับ นั่นคือ $n=500$ และ $\alpha=0.01$ เนื่องจากผู้หญิงที่ใช้เครื่องสำอางยี่ห้อนี้มีจำนวน 95 คน จากการสุ่มตัวอย่างมา 500 คน จะได้ว่า $\hat{p}=\dfrac{95}{500}=0.19$

<u>ขั้นที่ 1:</u> ตั้งสมมติฐาน

เนื่องจากบริษัทขายเครื่องสำอางยี่ห้อหนึ่งมีความเชื่อว่าผู้หญิงไทยใช้เครื่องสำอางยี่ห้อนี้อย่างน้อย 20% นั่นคือ $p\geq 0.2$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0: p\geq 0.2$$
$$H_1: p<0.2$$

<u>ขั้นที่ 2:</u> เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจาก $n\geq 30$ ดังนั้น จึงเลือกค่าสถิติ

$$Z=\frac{\hat{p}-p}{\sqrt{p(1-p)/n}}$$

<u>ขั้นที่ 3:</u> คำนวณค่าสถิติ

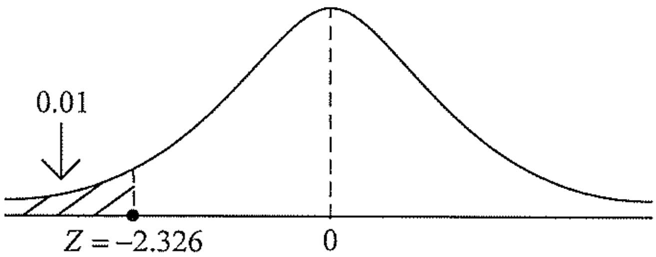
$$Z=\frac{\hat{p}-p}{\sqrt{p(1-p)/n}}=\frac{0.19-0.2}{\sqrt{(0.2)(1-0.2)/500}}=-0.56$$

<u>ขั้นที่ 4:</u> กำหนดระดับนัยสำคัญ

$$\alpha=0.01$$

<u>ขั้นที่ 5:</u> สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1: p<0.2$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านซ้ายของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha=0.01$ ดังรูป



โดยใช้ตารางที่ 3.2 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $Z < -2.326$

<u>ขั้นที่ 6:</u> สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($Z = -0.56$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($Z < -2.326$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.01

สรุปได้ว่า ผู้หญิงไทยใช้เครื่องสำอางยี่ห้อนี้อย่างน้อย 20% ที่ระดับนัยสำคัญ 0.01 นั่นคือ ความเชื่อของบริษัทนี้สมเหตุสมผลที่ระดับนัยสำคัญ 0.01 ■

<u>**ตัวอย่างที่ 7.3.2**</u> บริษัทผลิตเครื่องทำน้ำอุ่นแห่งหนึ่งกล่าวอ้างว่า 70% ของบ้านพักในจังหวัดเชียงใหม่มีเครื่องทำน้ำอุ่นที่บริษัทนี้ผลิตติดตั้งอยู่ ถ้าจากการสำรวจบ้านพักในจังหวัดเชียงใหม่จำนวน 150 หลัง พบว่ามีอยู่ทั้งหมด 80 หลัง ที่มีเครื่องทำน้ำอุ่นที่บริษัทนี้ผลิตติดตั้งอยู่ คำกล่าวอ้างของบริษัทผู้ผลิตเครื่องทำน้ำอุ่นแห่งนี้น่าเชื่อถือหรือไม่ที่ระดับนัยสำคัญ 10%

วิธีทำ กำหนดให้ p แทนค่าสัดส่วนของบ้านพักในจังหวัดเชียงใหม่ที่มีเครื่องทำน้ำอุ่นที่บริษัทนี้ผลิตติดตั้งอยู่ จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับนัยสำคัญมีค่าเท่ากับ 150 และ 0.1 ตามลำดับ นั่นคือ $n = 150$ และ $\alpha = 0.1$ เนื่องจากนี้มีบ้านพักในจังหวัดเชียงใหม่จำนวน 80 หลัง ที่มีเครื่องทำน้ำอุ่นที่บริษัทนี้ผลิตติดตั้งอยู่ จากการสุ่มตัวอย่างมา 150 หลัง จะได้ว่า $\hat{p} = \frac{80}{150} = \frac{8}{15}$

<u>ขั้นที่ 1:</u> ตั้งสมมติฐาน

เนื่องจากบริษัทผลิตเครื่องทำน้ำอุ่นแห่งหนึ่งกล่าวอ้างว่า 70% ของบ้านพักในจังหวัดเชียงใหม่มีเครื่องทำน้ำอุ่นที่บริษัทนี้ผลิตติดตั้งอยู่ นั่นคือ $p = 0.7$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0 : p = 0.7$$
$$H_1 : p \neq 0.7$$

<u>ขั้นที่ 2:</u> เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจาก $n \geq 30$ ดังนั้น จึงเลือกค่าสถิติ

$$Z = \frac{\hat{p} - p}{\sqrt{p(1-p)/n}}$$

<u>ขั้นที่ 3:</u> คำนวณค่าสถิติ

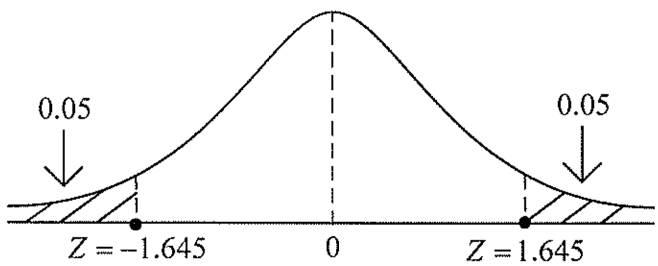
$$Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} = \frac{\frac{8}{15}-0.7}{\sqrt{(0.7)(1-0.7)/150}} = -4.45$$

<u>ขั้นที่ 4:</u> กำหนดระดับนัยสำคัญ

$$\alpha = 0.1$$

<u>ขั้นที่ 5:</u> สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : p \neq 0.7$ จึงทำให้เป็นการทดสอบสมมติฐานแบบสองข้าง โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านซ้ายและขวาของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha = 0.1$ ดังรูป



โดยใช้ตารางที่ 3.2 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $Z < -1.645$ หรือ $Z > 1.645$

<u>ขั้นที่ 6:</u> สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($Z = -4.45$) ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($Z < -1.645$) ดังนั้นจึงปฏิเสธ H_0 ที่ระดับนัยสำคัญ 0.1

จากข้อมูลที่ให้มาข้างต้น ไม่สามารถกล่าวได้ว่าค่าสัดส่วนของบ้านพักในจังหวัดเชียงใหม่ที่มีเครื่องทำน้ำอุ่นที่บริษัทนี้ผลิตติดตั้งอยู่มีค่าเท่ากับ 0.7 ที่ระดับนัยสำคัญ 0.1 นั่นคือ คำกล่าวอ้างของบริษัทผู้ผลิตเครื่องทำน้ำอุ่นแห่งนี้ไม่น่าเชื่อถือที่ระดับนัยสำคัญ 0.1 ■

<u>ตัวอย่างที่ 7.3.3</u> การเลือกตั้งประธานนักเรียนของโรงเรียนแห่งหนึ่ง มีผู้สมัครรับเลือกตั้งทั้งหมด 2 คน คือ นทีและ ธีระ ก่อนการเลือกตั้งมีนักเรียนกลุ่มหนึ่งได้คาดเดาเกี่ยวกับผลการเลือกตั้งไว้ว่า ธีระจะมีผู้เลือกเป็นประธานนักเรียนไม่เกิน 40% เพื่อสนับสนุนความเชื่อนี้นักเรียนกลุ่มนี้จึงได้ทำการสอบถามเกี่ยวกับความนิยมของผู้สมัครรับเลือกตั้งทั้งสองคนพบว่า จากการสอบถามนักเรียนจำนวน 100 คน มีผู้ที่ชื่นชอบนทีและจะเลือกนทีเป็นประธานนักเรียนจำนวน 58 คน ส่วนที่เหลือชื่นชอบธีระและจะเลือกธีระเป็นประธานนักเรียน การคาดเดาของนักเรียนกลุ่มนี้น่าเชื่อถือหรือไม่ ที่ระดับนัยสำคัญ 0.025

วิธีทำ กำหนดให้ p แทนค่าสัดส่วนของนักเรียนที่เลือกธีระเป็นประธานนักเรียน จากโจทย์ ขนาดของกลุ่มตัวอย่าง และระดับนัยสำคัญมีค่าเท่ากับ 100 และ 0.025 ตามลำดับ นั่นคือ $n = 100$ และ $\alpha = 0.025$ เนื่องจาก

ผู้ชื่นชอบนทีและจะเลือกนทีเป็นประธานนักเรียนมีจำนวน 70 คน จากการสุ่มตัวอย่างมา 100 คน จะได้ว่า

$$\hat{p} = \frac{100-58}{100} = 0.42$$

<u>ขั้นที่ 1:</u> ตั้งสมมติฐาน

เนื่องจากนักเรียนกลุ่มนี้คาดเดาเกี่ยวกับผลการเลือกตั้งไว้ว่า ธีระจะมีผู้เลือกเป็นประธานนักเรียนไม่เกิน 40% นั่นคือ $p \le 0.4$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0 : p \le 0.4$$
$$H_1 : p > 0.4$$

<u>ขั้นที่ 2:</u> เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจาก $n \ge 30$ ดังนั้น จึงเลือกค่าสถิติ

$$Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}}$$

<u>ขั้นที่ 3:</u> คำนวณค่าสถิติ

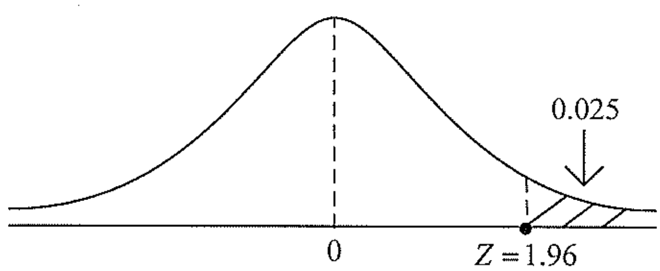
$$Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}} = \frac{0.42-0.4}{\sqrt{(0.4)(1-0.4)/100}} = 0.41$$

<u>ขั้นที่ 4:</u> กำหนดระดับนัยสำคัญ

$$\alpha = 0.025$$

<u>ขั้นที่ 5:</u> สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบปกติมาตรฐาน จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน (หรือโค้งปกติมาตรฐาน) ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : p > 0.4$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านขวาของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha = 0.025$ ดังรูป



โดยใช้ตารางที่ 3.2 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $Z > 1.96$

<u>ขั้นที่ 6:</u> สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($Z = 0.41$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($Z > 1.96$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.01

สรุปได้ว่า จำนวนนักเรียนที่เลือกธีระเป็นประธานนักเรียนมีไม่เกิน 40% ที่ระดับนัยสำคัญ 0.025 นั่นคือ การคาดเดาของนักเรียนกลุ่มนี้มีความน่าเชื่อถือ ที่ระดับนัยสำคัญ 0.025 ■

7.4 การทดสอบสมมติฐานเกี่ยวกับความแปรปรวนและส่วนเบี่ยงเบนมาตรฐานของประชากร

7.4.1 ค่าสถิติที่ใช้ในการทดสอบสมมติฐานเกี่ยวกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของประชากร

การทดสอบสมมติฐานเกี่ยวกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของประชากร

เมื่อประชากรมีการแจกแจงแบบปกติ ดังนั้น ค่าสถิติที่ใช้ในการทดสอบสมมติฐานคือ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

องศาอิสระ (degree of freedom) เท่ากับ $n-1$

เมื่อ σ^2 หมายถึง ความแปรปรวนของประชากร

s^2 หมายถึง ความแปรปรวนของกลุ่มตัวอย่าง

n หมายถึง ขนาดของกลุ่มตัวอย่าง

7.4.2 ขอบเขตปฏิเสธสมมติฐานเกี่ยวกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของประชากร

จากค่าสถิติที่ใช้ทดสอบสามารถสรุปขอบเขตปฏิเสธสมมติฐานเกี่ยวกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของประชากรได้ดังนี้

การทดสอบสมมติฐานเกี่ยวกับความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานของประชากร

1. การทดสอบแบบข้างเดียว

แบบที่ 1:

$$H_0 : \sigma \le \sigma_0 \quad \text{หรือ} \quad H_0 : \sigma^2 \le \sigma_0^2$$
$$H_1 : \sigma > \sigma_0 \qquad\qquad H_1 : \sigma^2 > \sigma_0^2$$

ขอบเขตปฏิเสธสมมติฐานคือ $\chi^2 > \chi_{\alpha}^2$

แบบที่ 2:

$$H_0 : \sigma \ge \sigma_0 \quad \text{หรือ} \quad H_0 : \sigma^2 \ge \sigma_0^2$$
$$H_1 : \sigma < \sigma_0 \qquad\qquad H_1 : \sigma^2 < \sigma_0^2$$

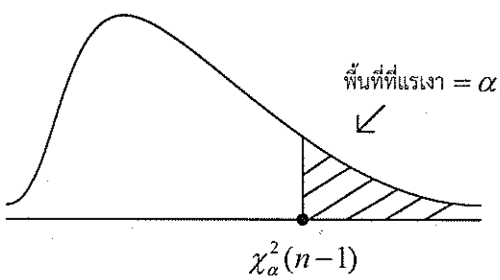
ขอบเขตปฏิเสธสมมติฐานคือ $\chi^2 < \chi_{1-\alpha}^2$

2. การทดสอบแบบสองข้าง

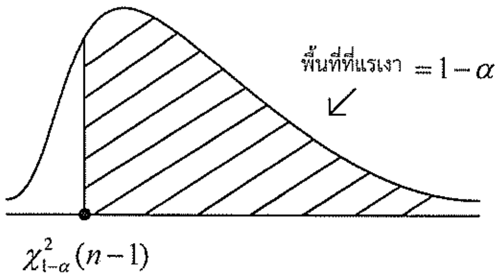
$$H_0 : \sigma = \sigma_0 \quad \text{หรือ} \quad H_0 : \sigma^2 = \sigma_0^2$$
$$H_1 : \sigma \ne \sigma_0 \qquad\qquad H_1 : \sigma^2 \ne \sigma_0^2$$

ขอบเขตปฏิเสธสมมติฐานคือ $\chi^2 < \chi_{1-\frac{\alpha}{2}}^2$ หรือ $\chi^2 > \chi_{\frac{\alpha}{2}}^2$

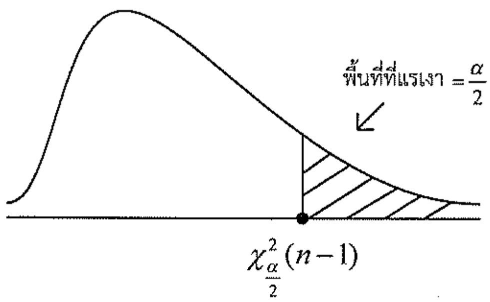
เมื่อ $\chi_{\alpha}^2(n-1)$ หมายถึง ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์



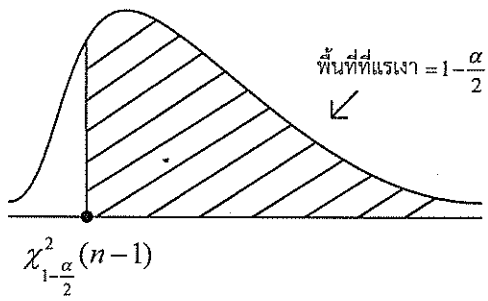
$\chi_{1-\alpha}^2(n-1)$ หมายถึง ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์



$\chi^2_{\frac{\alpha}{2}}(n-1)$ หมายถึง ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์



$\chi^2_{1-\frac{\alpha}{2}}(n-1)$ หมายถึง ตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบทีด้วยองศาอิสระเท่ากับ $n-1$ ซึ่งมีค่าเป็นไปตามรูปต่อไปนี้ เมื่อพิจารณากราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์



<u>ตัวอย่างที่ 7.4.1</u> ในกระบวนการเคลือบผิวแผ่น CD เราจะยอมรับว่ากระบวนการเคลือบผิวดังกล่าวได้มาตรฐาน ถ้าส่วนเบี่ยงเบนมาตรฐานของความหนาของสารเคลือบผิวที่วัดจากแผ่น CD มีค่าไม่เกิน 0.50 ม.ม. ได้มีการเลือกแผ่น CD มา 1 แผ่นแล้วทำตัดออกเป็น 15 ชิ้น คำนวณหาส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่างได้ 0.64 ม.ม. จงตัดสินว่าจะสามารถยอมรับกระบวนการผลิตดังกล่าวว่าได้มาตรฐานหรือไม่ ที่ระดับความเชื่อมั่น 95 % สมมติให้ความหนาของสารเคลือบผิวมีการแจกแจงแบบปกติ

วิธีทำ จากโจทย์ขนาดของกลุ่มตัวอย่าง ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง และระดับความเชื่อมั่นมีค่าเท่ากับ 15, 0.64 และ 0.05 ตามลำดับ นั่นคือ $n=15,\ s=0.64$ และ $\alpha=0.05$
<u>ขั้นที่ 1:</u> ตั้งสมมติฐาน

ในกระบวนการเคลือบผิวแผ่น CD จะยอมรับว่ากระบวนการเคลือบผิวดังกล่าวได้มาตรฐาน ถ้าส่วนเบี่ยงเบนมาตรฐานของความหนาของสารเคลือบผิวที่วัดจากแผ่น CD มีค่าไม่เกิน 0.50 ม.ม. นั่นคือ $\sigma \le 0.5$ ทำให้ตั้งสมมติฐานได้ดังนี้

$$H_0 : \sigma \le 0.5$$
$$H_1 : \sigma > 0.5$$

ขั้นที่ 2: เลือกค่าสถิติที่ใช้ในการทดสอบ

เนื่องจากความหนาของสารเคลือบผิวมีการแจกแจงแบบปกติ ดังนั้น จึงเลือกค่าสถิติ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

โดยที่ องศาอิสระเท่ากับ $n-1$

ขั้นที่ 3: คำนวณค่าสถิติ

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(15-1)(0.64)^2}{(0.5)^2} = 22.9376$$

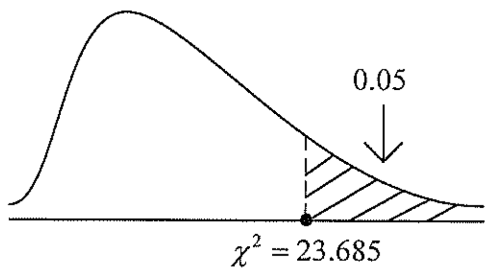
โดยที่ องศาอิสระเท่ากับ $n-1 = 15-1 = 14$

ขั้นที่ 4: กำหนดระดับนัยสำคัญ

$$\alpha = 0.05$$

ขั้นที่ 5: สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : \sigma > 0.5$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านขวาของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha = 0.05$ ดังรูป



โดยใช้ตารางที่ 4.2 องศาอิสระเท่ากับ 14 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $\chi^2 > 23.685$

ขั้นที่ 6: สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($\chi^2 = 22.9376$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($\chi^2 > 23.685$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.05

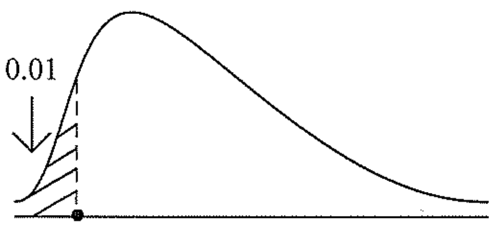
สรุปได้ว่า ส่วนเบี่ยงเบนมาตรฐานของสารเคลือบผิวมีค่าไม่เกิน 0.50 มิลลิเมตร ที่ระดับนัยสำคัญ 0.05 นั่นคือ กระบวนการผลิตดังกล่าวได้มาตรฐานที่ระดับนัยสำคัญ 0.05 ■

<u>ขั้นที่ 4:</u> กำหนดระดับนัยสำคัญ

$$\alpha = 0.01$$

<u>ขั้นที่ 5:</u> สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : \sigma^2 < 2.25$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านซ้ายของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha = 0.01$ ดังรูป



$\chi^2 = 7.633$

โดยใช้ตารางที่ 4.1 องศาอิสระเท่ากับ 19 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $\chi^2 < 7.633$

<u>ขั้นที่ 6:</u> สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($\chi^2 = 16.7884$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($\chi^2 < 7.633$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.01

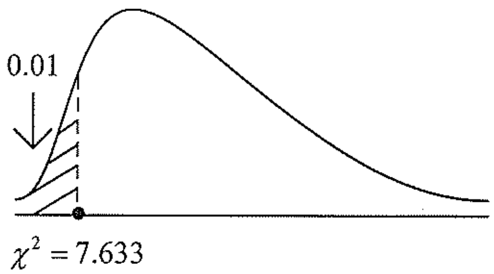
จากข้อมูลที่ให้มาข้างต้น ไม่สามารถกล่าวได้ว่าความแปรปรวนของความยาวชิ้นส่วนเครื่องยนต์ที่ผลิตมีค่าน้อยกว่า 2.25 มิลลิเมตร ที่ระดับนัยสำคัญ 0.01 นั่นคือ ความเชื่อของบริษัทผู้ผลิตไม่สมเหตุสมผลที่ระดับนัยสำคัญ 0.01 ■

<u>ขั้นที่ 4:</u> กำหนดระดับนัยสำคัญ

$$\alpha = 0.01$$

<u>ขั้นที่ 5:</u> สร้างขอบเขตปฏิเสธสมมติฐาน

เนื่องจากค่าสถิติที่เลือกในขั้นที่ 3 เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงความน่าจะเป็นแบบไคสแควร์ จึงใช้กราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ ในการพิจารณาขอบเขตปฏิเสธสมมติฐาน และเนื่องจากสมมติฐาน $H_1 : \sigma^2 < 2.25$ จึงทำให้เป็นการทดสอบสมมติฐานแบบข้างเดียว โดยมีขอบเขตปฏิเสธสมมติฐานอยู่ด้านซ้ายของกราฟ และมีพื้นที่ของขอบเขตปฏิเสธสมมติฐานเท่ากับระดับนัยสำคัญ $\alpha = 0.01$ ดังรูป



โดยใช้ตารางที่ 4.1 องศาอิสระเท่ากับ 19 และสมบัติของกราฟฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ จะได้ว่าขอบเขตปฏิเสธสมมติฐานคือ $\chi^2 < 7.633$

<u>ขั้นที่ 6:</u> สรุปผลการทดสอบ

เนื่องจากค่าสถิติที่คำนวณได้ในขั้นที่ 3 ($\chi^2 = 16.7884$) ไม่ตกอยู่ในขอบเขตปฏิเสธสมมติฐานที่หาได้จากขั้นที่ 5 ($\chi^2 < 7.633$) ดังนั้นจึงยอมรับ H_0 ที่ระดับนัยสำคัญ 0.01

จากข้อมูลที่ให้มาข้างต้น ไม่สามารถกล่าวได้ว่าความแปรปรวนของความยาวชิ้นส่วนเครื่องยนต์ที่ผลิตมีค่าน้อยกว่า 2.25 มิลลิเมตร ที่ระดับนัยสำคัญ 0.01 นั่นคือ ความเชื่อของบริษัทผู้ผลิตไม่สมเหตุสมผลที่ระดับนัยสำคัญ 0.01 ■