สถิติเพื่อการตัดสินใจ

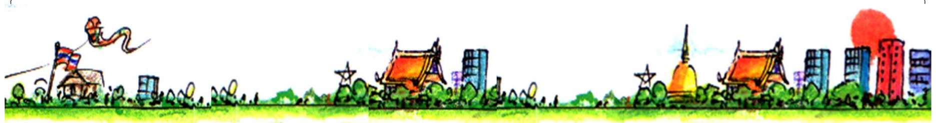
รองศาสตราจารย์พิษณุ เจียวคุณ
ภาควิชาสถิติ คณะวิทยาศาสตร์

สอบกลางภาค วันอังคารที่ 3 มีนาคม 2558 เวลา 8.00-11.00 น.

สถิติสำคัญอย่างไร

✓ สถิติเป็นศาสตร์ที่ว่าด้วยการจัดการข้อมูล ประกอบด้วย การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการแปลความหมายของข้อมูล

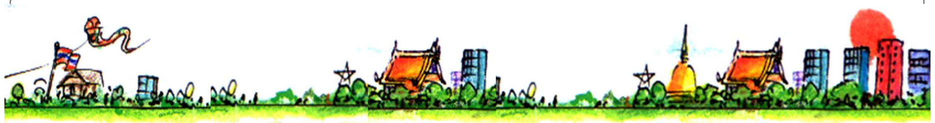
✓ เป็นเครื่องมือที่ใช้ในการตัดสินใจ โดยใช้ความน่าจะเป็น



สถิติสำคัญอย่างไร

✓ สถิติเป็นศาสตร์ที่ว่าด้วยการจัดการข้อมูล ประกอบด้วย การเก็บรวบรวมข้อมูล การนำเสนอข้อมูล การวิเคราะห์ข้อมูล และการแปลความหมายของข้อมูล

✓ เป็นเครื่องมือที่ใช้ในการตัดสินใจ โดยใช้ความน่าจะเป็น



ความน่าจะเป็น (Probability)

คือดัชนีชี้วัดความเป็นไปได้ของเหตุการณ์ ซึ่งเป็นองค์ประกอบหนึ่งของการตัดสินใจ ภายใต้ความไม่แน่นอน (Uncertainly) โดยมีองค์ประกอบที่สำคัญคือ การทดลองสุ่ม ปริภูมิตัวอย่าง และเหตุการณ์

การทดลองสุ่ม (Random Experiment)

เป็นปรากฏการณ์ใด ๆ ที่เราไม่ทราบผลการทดลองจนกว่าการทดลองนั้นจะสิ้นสุดลง แต่เราจะทราบถึงขอบเขตของผลลัพธ์ที่เกิดขึ้น เช่น การทดลองโยนเหรียญ การรักษาผู้ป่วยเป็นต้น

ปริภูมิตัวอย่าง (Sample space): S

คือเซตของผลลัพธ์ที่เป็นไปได้ทั้งหมดในการทดลองสุ่มใด ๆ เช่น การทดลองโยนเหรียญที่เที่ยงตรงจำนวน 2 เหรียญ ดังนั้น

$$S = \{ HH, HT, TH, TT \}$$

เหตุการณ์ (Event) : E

คือเซตย่อยของปริภูมิตัวอย่าง หรือเป็นกลุ่มของผลลัพธ์ที่สนใจจากปริภูมิตัวอย่าง จำนวนสมาชิกในเหตุการณ์จะต้องไม่เกินจำนวนสมาชิกในปริภูมิตัวอย่าง

ประเภทของความน่าจะเป็น

1. Classical Probability
2. Frequency Probability
3. Subjective Probability

ความน่าจะเป็นแบบ Classical Probability

พิจารณาจำนวนสมาชิกของเหตุการณ์เทียบกับจำนวนสมาชิกที่เป็นไปได้ทั้งหมด ดังนั้น

$$p(E) = \frac{n(E)}{n(S)}$$

ตัวอย่าง

โยนเหรียญที่เที่ยงตรง 2 เหรียญ จงหาความน่าจะเป็นที่เหรียญจะหงายหน้าก้อย 1 เหรียญและหงายหน้าหัว 1 เหรียญ

วิธีทำ

ปริภูมิตัวอย่างของการทดลองสุ่มนี้คือ

S= {HH,HT,TH,TT} ดังนั้น n(S)= 4

เหตุการณ์ที่สนใจคือ E ={HT,TH}

ดังนั้น n(E)=2

P(E) = n(E)/n(S) =2/4 =0.5

ตัวอย่าง

กล่องใบหนึ่งบรรจุลูกบอลขนาดเดียวกัน จำนวน 10 ลูก ประกอบด้วยสีแดง 1 ลูก สีเขียว 2 ลูก สีเหลือง 3 ลูก และสีน้ำเงิน 4 ลูก สุ่มลูกบอล 2 ลูก จงหาความน่าจะเป็นที่จะได้ลูกบอลดังต่อไปนี้

ตัวอย่าง

ก. ความน่าจะเป็นที่จะได้ลูกบอลสีเดียวกัน

ข. ความน่าจะเป็นที่จะได้ลูกบอลสีเขียว 1 ลูก

ค. ความน่าจะเป็นที่จะได้ลูกบอลสีเขียว 1 ลูก และสีแดง 1 ลูก

วิธีทำ

หาจำนวนวิธีในปริภูมิตัวอย่าง n(S) = $C_{10,2} = 45$

ก. จำนวนวิธีที่จะได้ลูกบอลสีเหมือนกัน

ลูกบอลสีเขียวเหมือนกัน $C_{2,2} = 1$

ลูกบอลสีเหลืองเหมือนกัน $C_{3,2} = 3$

ลูกบอลสีน้ำเงินเหมือนกัน $C_{4,2} = 6$

ตัวอย่าง

ดังนั้น $$p(E) = \frac{n(E)}{n(S)} = \frac{10}{45} = 0.22$$

ข. ความน่าจะเป็นที่จะได้ลูกบอลสีเขียว 1 ลูก

จำนวนวิธี = $C_{2,1}.C_{8,1} = 16$

$$p(E) = \frac{n(E)}{n(S)} = \frac{16}{45} = 0.36$$

ตัวอย่าง

ค. ความน่าจะเป็นที่จะได้ลูกบอลสีเขียว 1 ลูก และสีแดง 1 ลูกจำนวนวิธี = $C_{1,1}.C_{2,1} = 2$

$$p(E) = \frac{n(E)}{n(S)} = \frac{2}{45} = 0.04$$

ความน่าจะเป็นแบบ Frequency Probability

เป็นการกำหนดความน่าจะเป็นโดยอ้างอิงความถี่ของการเกิดผลลัพธ์ที่สนใจ โดยทำการทดลองหลายๆ ครั้ง จนกระทั่งสัดส่วนของการเกิดเหตุการณ์ที่สนใจมีค่าคงที่ ซึ่งค่าคงที่นี้คือค่า<u>ความน่าจะเป็น</u>

ความน่าจะเป็นแบบ Subjective Probability

เป็นการกำหนดความน่าจะเป็นโดยอาศัยประสบการณ์ของผู้เชี่ยวชาญ ยิ่งมีข้อมูลมาก การหาความน่าจะเป็นจะมีความแม่นยำมากขึ้น เช่น การทำนายผลตอบแทนการลงทุน หุ้น ฯลฯ

ค่าคาดหมาย (Expectation)

ค่าเฉลี่ยของการเกิดขึ้นของค่าของตัวแปรที่สนใจ ภายใต้ความน่าจะเป็นที่ตัวแปรแต่ละค่าจะมีโอกาสเกิดขึ้น

ค่าคาดหมาย (Expectation) : E(X)

นิยาม กำหนดให้ X เป็นตัวแปร ประกอบด้วยค่าของตัวแปร X1,X2, ...,Xn ด้วยความน่าจะเป็น P1, P2,.....,Pn ตามลำดับ ดังนั้นค่าคาดหมาย คือ

$$E(X) = \sum_{i=1}^{n} x_i p_i$$

ตัวอย่าง

การศึกษาผลตอบแทนของการถูกรางวัลสลากกินแบ่งรัฐบาล เลขท้าย 2 ตัวจำนวน 1 ใบ ถ้าถูกรางวัลจะได้เงินรางวัล 2,000 บาท ถ้าไม่ถูกรางวัลต้องจ่ายเงินค่าสลาก 80 บาท

ตัวอย่าง

ให้ X แทนผลตอบแทน ดังนั้น

X = -80 , 2,000

P(X=-80) = 99/100

P(X=2,000) = 1/100

E(X) = -80(99/100)+2,000(1/100)=-59.2

ตัวอย่าง

ค่าคาดหมายของผลตอบแทนมีค่าเท่ากับ -59.20 บาท หมายความว่า ถ้าเราซื้อสลากกินแบ่งรัฐบาลจำนวน 1 ใบ โดยที่เราสนใจเฉพาะรางวัลเลขท้าย 2 ตัว คาดว่าจะเสียเงินโดยเฉลี่ย 59.20 บาทต่อใบ

ตัวอย่าง

ให้ X แทนผลตอบแทน ถ้าเราต้องจ่ายเงินค่าสลากก่อน ดังนั้น

X = -80 , 1,920

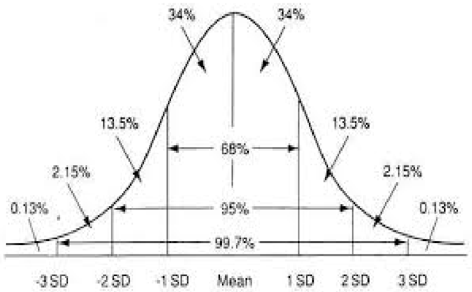
P(X=-80) = 99/100

P(X=1,920) = 1/100

การแจกแจงปกติ

โดยทั่วไปถ้ามีข้อมูลจำนวนมาก ถ้านำมาจัดเรียงแล้วจะพบว่า ข้อมูลที่มีค่ามากหรือน้อยจะมีจำนวนไม่มาก ส่วนข้อมูลที่มีค่าปานกลาง จะมีจำนวนมาก เมื่อนำมาพล๊อตกราฟแล้วจะพบว่า จะมีลักษณะคล้ายระฆังคว่ำ

รูปการแจกแจงปกติ



การแจกแจงปกติ

ค่าเฉลี่ยของข้อมูลจะอยู่บริเวณตรงกลางของรูปโค้งระฆังคว่ำ รูปการแจกแจงปกติจะมีลักษณะสมมาตร (Symmetry curve)

การแจกแจงปกติมาตรฐาน

(Standard normal distribution)

ถ้าให้ X แทนตัวแปรสุ่มที่มีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ μ และมีค่าเบี่ยงเบนมาตรฐานเท่ากับ σ แล้ว เราสามารถแปลงค่า X ให้เป็นค่า Z ซึ่งเป็นการแจกแจงปกติมาตรฐาน โดยที่

การแจกแจงปกติมาตรฐาน

(Standard normal distribution)

โดยที่

$$Z = \frac{x-\mu}{\sigma}$$

Z มีค่าเฉลี่ยเท่ากับ 0 และมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 1 หรือ $Z \sim N(0,1)$

การแจกแจงปกติมาตรฐาน

(Standard normal distribution)

การหาค่าความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงปกติมาตรฐาน Z สามารถเปิดตารางการแจกแจงปกติเพื่อหาพื้นที่ใต้โค้งได้

การแจกแจงปกติมาตรฐาน

(Standard normal distribution)

ตัวอย่าง

$P(0<Z<1.23) = 0.3907$

$P(Z>1.23) = 0.5-0.3907= 0.1093$

$P(Z<-1.23) = 0.5-0.3907= 0.1093$

การแจกแจงปกติมาตรฐาน
(Standard normal distribution)

ตัวอย่าง

$P(-2.5<Z<-1.5) = 0.4938-0.4332=0.0606$

$P(Z>c) = 0.025$, $c = 1.96$

ตัวอย่าง

การสอบครั้งหนึ่งมีนักศึกษาเข้าสอบทั้งหมด 800 คน คะแนนสอบมีค่าเฉลี่ยเท่ากับ 60 คน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 10 คะแนน

ก. สุ่ม นศ.จำนวน 1 คน จงหาความน่าจะเป็นที่จะได้ นศ.ที่มีคะแนนมากกว่า 50 คะแนน

ตัวอย่าง

ให้ X เป็นคะแนนสอบของนักศึกษา

$\mu=60$ $\sigma=10$

หา $P(X>50) = P(Z>-1)$

$= 0.3413+0.5$

$= 0.8413$

ตัวอย่าง

ข.ถ้าอาจารย์ผู้สอนกำหนดเกรด F ถ้าได้คะแนนน้อยกว่า 45 คะแนน คาดว่าจะมี นศ.ที่สอบได้เกรด F กี่คน

$P(X<45) = P(Z<-1.5) = 0.5-0.4332$

$= 0.0668$

ตัวอย่าง

นั่นคือมี นศ.คาดว่าจะสอบได้เกรด F ด้วยความน่าจะเป็นเท่ากับ 0.0668 หรือคิดเป็น 6.68%

ถ้ามีนักศึกษา 800 คน คาดว่าจะมีผู้ที่สอบได้เกรด F จำนวน 0.0668(800) = 53.44 หรือประมาณ 54 คน

ตัวอย่าง

ค.ถ้ามีนักศึกษาได้รับเกรด A จำนวน 80 คน ผู้ที่ได้เกรด A ต้องได้คะแนนอย่างน้อยเท่าไร

แนวคิด คือต้องหาได้ผู้ที่ได้เกรด A คิดเป็นสัดส่วน 80/800 = 0.10 หรือความน่าจะเป็นที่จะได้เกรด A เท่ากับ 0.10

ตัวอย่าง

เมื่อเทียบกับตาราง Z จะพบว่า

$P(0<Z<1.28) \approx 0.40$

ดังนั้น $\frac{a-60}{10} = 1.28$ a = 72.8

นักศึกษาที่ได้รับเกรด A จะต้องได้คะแนนตั้งแต่ 72.8 ขึ้นไป จึงจะได้เกรด A

ตัวอย่าง

ถ้าอายุการใช้งานของเครื่องคำนวณชนิดหนึ่งมีการแจกแจงปกติ มีค่าเฉลี่ย 48 เดือน ส่วนเบี่ยงเบนมาตรฐาน 12 เดือน บริษัทรับประกันถ้าเครื่องคำนวณเสียภายใน 18 เดือน จะจ่ายเงินชดเชยเครื่องละ 2000 บาท ถ้าบริษัทผลิตสินค้าออกมา 1500 เครื่อง บริษัทต้องเตรียมเงินชดเชยให้ลูกค้าเท่าไร ?

ตัวอย่าง

ถ้าให้ X แทนอายุการใช้งานของเครื่องคำนวณ

หา P(X<18)

ดังนั้น P(X<18) = P(Z<-2.5)=0.0062

ดังนั้นเครื่องคำนวณจะเสียทั้งหมด = 1500x 0.0062

คิดเป็น 9.3 หรือประมาณ 9 เครื่อง คิดเป็นเงิน 9x2000

เท่ากับ 18000 บาท