บทที่ 1
ความน่าจะเป็น

คำนำ

ในปัญหาการตัดสินใจภายใต้ความไม่แน่นอน มีอยู่สิ่งหนึ่งที่จะต้องพบกับการตัดสินใจขึ้นอยู่กับปรากฏการณ์ซึ่งมีความไม่แน่นอน ความไม่แน่นอนนี้เป็นสาเหตุทำให้ปัจจัยที่ควบคุมปรากฏการณ์ธรรมชาติได้ไม่มากนัก สิ่งหนึ่งที่จะทำได้คือนำเข้าไปสู่แบบจำลองทางคณิตศาสตร์และดำเนินแบบเชิงปริมาณ โดยทั่ว ๆ ไปถ้าปรากฏการณ์ธรรมชาติแสดงออกมาในลักษณะที่ค่อนข้างจะแน่นอนเพื่อว่าการกระจายสามารถอธิบายได้ด้วยแบบจำลองความน่าจะเป็น ในหัวข้อนี้ก็จะกล่าวในลักษณะของแบบจำลองความน่าจะเป็น

Sample Space

สมมติว่าอุปสงค์สำหรับผลิตภัณฑ์ตลอดระยะเวลาหนึ่ง และอุปสงค์ก็ไม่คงที่แต่แสดงออกเป็นรูปต่าง ๆ กันในระยะเวลาหนึ่งเดือน ผลลัพธ์ของอุปสงค์ไม่สามารถทำนายได้แน่นอน แต่ละผลลัพธ์ที่เป็นไปได้สามารถอธิบายได้เป็นค่าใดค่าหนึ่งของ 0, 1, 2,... ซึ่งเป็นเซทของเลขจำนวนเต็มบวก เซทของผลลัพธ์ของอุปสงค์เหล่านี้เรียกว่า sample space แต่ละผลลัพธ์ใน sample space เรียกว่าจุดหนึ่ง อุปสงค์ที่เป็นไปได้อาจประกอบด้วย N ในเมื่อ N ใช้แทนขนาดของประชากร ดังนั้น sample space ประกอบด้วยเซทของเลขจำนวนเต็ม 0, 1, 2,...N ตัวอย่างของการทดลองอื่น ๆ อาจเกี่ยวกับเวลาจนกระทั่งลูกค้าคนแรกเริ่มมาถึงร้าน เนื่องจากว่าลูกค้าอาจมาถึงร้านที่เวลาใดเวลาหนึ่งจนกระทั่งร้านปิด สำหรับการทดลองนี้ sample space สามารถพิจารณาให้เห็นจุดทั้งหมดบนเส้นระหว่างศูนย์ถึงแปดชั่วโมง sample space ประกอบด้วยจุดทั้งหมดอย่างเช่น $0 \leqslant X \leqslant 8$

สมมติว่าเราดัดแปลงตัวอย่างแรกเกี่ยวกับอุปสงค์ระหว่างสองเดือนแรก sample space ประกอบด้วยจุดทั้งหมด (X_1, X_2) ในเมื่อ X_1 ใช้แทนอุปสงค์ระหว่างเดือนแรก $X_1 = 0, 1, 2, \ldots$ และ X_2 ใช้แทนอุปสงค์ระหว่างเดือนที่สอง $X_2 = 0, 1, 2, \ldots$ ดังนั้น sample space ประกอบ

ด้วยเซทของจุดทั้งหมดที่เป็นไปได้ w ในเมื่อ w ใช้แทนคู่หนึ่งของมูลค่าจำนวนเต็มบวก (X_1, X_2) จุด w = (3, 6) ใช้แทนผลลัพธ์ที่เป็นไปได้ของการทดลองในเมื่ออุปสงค์ในเดือนแรกเป็น 3 หน่วยและอุปสงค์ในเดือนที่สองเป็น 6 หน่วย ในลักษณะเดียวกัน การทดลองสามารถขยายออกเป็นอุปสงค์ระหว่าง n เดือนแรกในสภาวะนี้ sample space ประกอบด้วยจุดทั้งหมดที่เป็นไปได้ $w = (X_1, X_2, \ldots, X_n)$ ในเมื่อ X_i ใช้แทนอุปสงค์ระหว่างเดือนที่ i

การทดลองที่เกี่ยวกับเวลาจนกระทั่งคนแรกได้มาถึงร้านสามารถดัดแปลงได้ สมมติว่าการทดลองเป็นการวัดเวลาของลูกค้าคนแรกมาถึงร้านของแต่ละครั้งในสองวัน เซทของผลลัพธ์ทั้งหมดที่เป็นไปได้ของการทดลองประกอบด้วยจุดทั้งหมด (X_1, X_2), $0 \leqslant X_1, X_2 \leqslant 8$ ในเมื่อ X_1 ใช้แทนเวลาลูกค้าคนแรกมาถึงร้านของวันแรกและ X_2 ใช้แทนเวลาลูกค้าคนแรกมาถึงร้านในวันที่สอง ดังนั้น sample space ของเซทของจุดทั้งหมดที่เป็นไปได้ w ในเมื่อ w ใช้แทนจุดในสองมิติ

การทดลองนี้สามารถขยายเวลามาถึงร้านของลูกค้าคนแรกของแต่ละครั้งของ n วัน sample space ประกอบด้วยจุดทั้งหมด $w = (X_1, X_2, \ldots, X_n)$ ดังเช่น $0 \leqslant X_i \leqslant 8$ $(i = 1, 2, \ldots, n)$ ในเมื่อ X_i ใช้แทนเวลาลูกค้ามาถึงร้านคนแรกของวันที่ i •

ตัวแปรเชิงสุ่ม

สมมติว่าการทดลองหนึ่งเกี่ยวกับการวัดเวลาของลูกค้ามาถึงร้านคนแรกในช่วงสองวันแรก เพื่อกำหนดว่าเวลาอะไรร้านจึงจะเปิด การทดลองก่อน ๆ เจ้าของร้านตัดสินใจว่า ถ้าหากว่าเวลาเฉลี่ยมากกว่าหนึ่งชั่วโมงหลังจากนั้นเขาจะไม่เปิดร้านของเขาจนกระทั่ง 10.00 น. (9.00 น. เป็นเวลาที่เปิดครั้งก่อน) ค่าเฉลี่ยของ X_1 กับ X_2 (เวลาที่มาถึงร้านสองครั้งไม่ได้เป็นจุดใน sample space ดังนั้นเขาไม่สามารถทำการตัดสินใจโดยพิจารณาที่ผลลัพธ์การทดลองของเขา เขาทำการตัดสินใจตามผลลัพธ์ของกฎเฉลี่ย X_1 กับ X_2 ของแต่ละจุด (X_1, X_2) ใน sample space เซทนี้ ถูกแบ่งออกเป็นสองส่วน จุดเหล่านั้นซึ่งอยู่ต่ำกว่า 1 กับจุดเหล่านั้นซึ่งอยู่สูงกว่า 1 ถ้าผลที่ได้สังเกตของกฎนี้วางอยู่ในส่วนที่จุดมีค่ามากกว่า 1 ร้านจะเปิด 10.00 น. ในทางตรงข้ามร้านจะเปิด 9.00 น. ต่อไป กฎซึ่งกำหนดส่วนเฉลี่ยของ X_1 กับ X_2 ของแต่ละจุดใน sample space เรียกว่าตัวแปรเชิงสุ่ม ดังนั้นตัวแปรเชิงสุ่มเป็นฟังก์ชันของค่าตัวเลขที่ได้กำหนดตลอด sample space ตัวแปรเชิงสุ่มใช้ตัวอักษรตัวโตและมูลค่าตัวแปรเชิงสุ่มมีอักษรตัวอื่นประกอบด้วย อย่างเช่น $\bar{X}$ ควรเขียนเป็น $\bar{X}(w)$ ดังตัวอย่าง $\bar{X}(1, 2) = (1+2)/2 = 1.5$, $\bar{X}(1.6, 1.8) =$

$(1.6+1.8)/2 = 1.7$ มูลค่าที่ตัวแปรเชิงสุ่ม $\bar{X}$ ได้มาจากเซทของค่า $\bar{X}$ อย่างเช่น $0 \leqslant \bar{X} \leqslant 8$ ตัวแปรเชิงสุ่มอื่น ๆ $\bar{X}_1$ สามารถอธิบายได้ดังนี้ แต่ละ w ใน sample space ตัวแปรเชิงสุ่มที่ไม่เกี่ยวกับ X_2 Coordinate และแปลง X_1 Coordinate เป็นตัวมันเอง ตัวแปรเชิงสุ่มนี้ใช้แทนเวลามาถึงของลูกค้าคนแรกของวันแรก ดังตัวอย่าง $\bar{X}_1(1, 2) = 1$, $\bar{X}_1(1.6, 1.8) = 1.6$ $\bar{X}_1(1.5, 1.5) = 1.5$, $\bar{X}_1(8, 8) = 8$ มูลค่าตัวแปรเชิงสุ่ม X_1 ได้มาจากเซทของค่า X_1 อย่างเช่น $0 \leqslant X_1 \leqslant 8$ ในทำนองเดียวกันตัวแปรเชิงสุ่ม X_2 สามารถอธิบายได้โดยใช้เวลามาถึงของลูกค้าคนแรกในวันที่สอง ตัวแปรเชิงสุ่มยังหาได้จากวิธีการอื่น ๆ ซึ่งจะไม่กล่าวในที่นี้

ความน่าจะเป็นและการแจกแจงความน่าจะเป็น

เรากลับไปยังตัวอย่างของอุปสงค์สำหรับผลิตภัณฑ์ระหว่างหนึ่งเดือน สังเกตว่า อุปสงค์จริง ๆ ไม่ได้คงที่แต่แสดงออกของการกระจาย โดยเฉพาะการกระจายนี้สามารถอธิบายในรูปของความน่าจะเป็นที่กำหนดกับเหตุการณ์ทั้งหมดใน sample space ดังตัวอย่าง ถ้าเหตุการณ์เป็นเซท E_1 ในเมื่อ $E_1 = \{w = 0, w = 1, w = 2, \ldots w = 10\}$ แล้วเป็นที่เข้าใจว่า $P(E_1)$ เป็นความน่าจะเป็นของอุปสงค์ 10 หน่วย หรือน้อยกว่ารายละเอียดบางอย่างที่เป็นประโยชน์เกี่ยวกับอุปสงค์ที่คาดหวังจะเกิดขึ้น ถ้าหากว่าทราบ $P(E)$ สำหรับเซท E ทั้งหมดใน sample space แล้วเพื่อที่จะนิยาม Concept ของความน่าจะเป็นเป็นสิ่งที่เกินขอบเขตของหนังสือเล่มนี้อย่างไรก็ตาม คุณสมบัติของความน่าจะเป็นสามารถอธิบายได้

(1) $0 \leqslant P(E) \leqslant 1$ หมายความว่าความน่าจะเป็นของเหตุการณ์หนึ่งจะเป็นบวกเสมอและไม่เกินหนึ่ง

(2) ถ้า E_0 เป็นเหตุการณ์หนึ่งซึ่งไม่สามารถเกิดขึ้นใน sample space อย่างเช่น อุปสงค์ของ 5 หน่วยแล้ว $P(E_0) = 0$

(3) $P(S) = 1$ ถ้าเหตุการณ์เป็น sample space ทั้งหมด

(4) ถ้า E_1 กับ E_2 เป็นเหตุการณ์ไม่ร่วมกัน (mutually exclusive) ใน sample space แล้ว $P(E_1 + E_2) = P(E_1) + P(E_2)$ ในเมื่อ $(E_1 + E_2)$ เป็นเหตุการณ์ E_1 หรือ E_2 ดังนั้น ถ้า E_1 เป็นเหตุการณ์ของอุปสงค์ของ 0 หรือ 1 และ E_2 เป็นเหตุการณ์ของอุปสงค์ 4 หรือ 5 แล้ว ความน่าจะเป็นของอุปสงค์ 0, 1, 4 หรือ 5 อย่างเช่น $(E_1 + E_2)$ กำหนดได้โดย $P(E_1) + P(E_2)$ คุณสมบัติข้อนี้ถ้าจะกล่าวให้รัดกุมยิ่งขึ้น ดังนี้ ให้ n เป็นจำนวนของครั้งของการทดลองหนึ่ง m เป็นจำนวนของความสำเร็จที่เกิดขึ้นของเหตุการณ์ E ในจำนวนของ n ครั้งแล้ว $P(E)$ สามารถนิยามได้เป็น

$$P(E) = \lim_{n \to \infty} \frac{m}{n}$$

สมมติ limit หาค่าได้สำหรับปรากฏการณ์นั้น อัตรา m/n ตรงกับเงื่อนไขสำหรับความน่าจะเป็น

(1) $0 \leqslant m/n \leqslant 1$

(2) $0/n = 0$ (ถ้าเหตุการณ์ E ไม่สามารถเกิดขึ้นแล้ว $m = 0$)

(3) $n/n = 1$ (ถ้าเหตุการณ์ E ต้องเกิดขึ้นทุกครั้งของการทดลองแล้ว $m = n$)

(4) $(m_1 + m_2) / n = m_1/n + m_2/n$ ถ้า E_1 กับ E_2 เป็นเหตุการณ์ที่ไม่ร่วมกัน

ถ้าคุณสมบัติเหล่านี้เป็นจริงสำหรับ n มีค่าจำกัดแล้วเราก็ควรคาดหวังว่าเป็นจริงสำหรับ

$$P(E) = \lim_{n \to \infty} \frac{m}{n}$$

เราเห็นว่า เราสามารถสมมติสิ่งที่เป็นจริงของความน่าจะเป็น กำหนดกับเหตุการณ์ E ใน sample space อย่างไรก็ตาม เราสนใจตัวแปรเชิงสุ่มและการคำนวณหาความสัมพันธ์ระหว่างความน่าจะเป็นที่สมนัยกับเหตุการณ์ใน sample space กับความน่าจะเป็นที่สมนัยกับตัวแปรเชิงสุ่ม

ฟังก์ชั่นการแจกแจง

ฟังก์ชันการแจกแจงมีความสัมพันธ์กับตัวแปรเชิงสุ่ม เพื่อที่จะนิยามฟังก์ชันการแจกแจง จำเป็นจะต้องแนะนำสัญลักษณ์บางอย่าง ให้สัญลักษณ์ $E_b = \{X(w) \leqslant b\}$ เป็นเซทของผลลัพธ์ w ใน sample space ที่ก่อร่างเป็นเหตุการณ์ E_b อย่างเช่นตัวแปรเชิงสุ่ม X เป็นค่าน้อยกว่าหรือเท่ากับ b แล้ว $P(E_b)$ ก็เป็นความน่าจะเป็นของเหตุการณ์นี้ ดังตัวอย่าง สมมติการทดลองเป็นการวัดอุปสงค์สำหรับผลิตภัณฑ์ระหว่างหนึ่งเดือน ให้ $N = 99$ และสมมติว่าเหตุการณ์ $\{0\}, \{1\}, \{2\}, \ldots \{99\}$ แต่ละเหตุการณ์มีความน่าจะเป็น 1/100 ดังเช่น $P(0) = P(1) = P(2) = \ldots = P(99) = \frac{1}{100}$ ให้ตัวแปรเชิงสุ่ม X เป็นกำลังสองของอุปสงค์ ดังนั้น $\{X \leqslant 150\}$ เป็นเซท $E_b = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ (เนื่องจากว่ากำลังสองของเลขจำนวนเหล่านี้น้อยกว่า 150) ดังนั้น

$$P(E_b) = \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100}$$

$$\frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{13}{100}$$

ดังนั้น $P(E_b) = P\{X \leqslant b\} = \frac{13}{100}$ สำหรับตัวแปรเชิงสุ่ม X ที่กำหนดให้ $P\{X \leqslant b\}$ แสดงได้เป็น $F_X(b)$ เรียกว่าฟังก์ชัน การแจกแจง (C.D.F.) ของตัวแปรเชิงสุ่ม X และถูกกำหนดสำหรับมูลค่าจริงทั้งหมดของ b C.D.F. ของตัวแปรเชิงสุ่มเป็นฟังก์ชันของค่าตัวเลขถูกกำหนดสำหรับ b ทั้งหมด $-\infty \leqslant b \leqslant \infty$ ซึ่งมีคุณสมบัติดังนี้

(1) $F_x(b)$ เป็น nondecreasing function ของ b

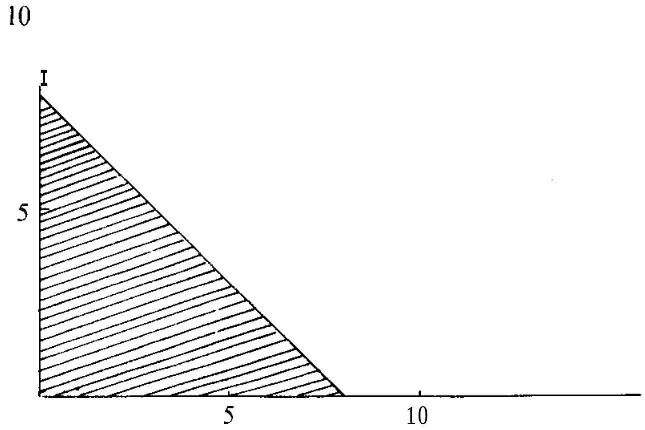
(2) $\lim_{b \to -\infty} F_x(b) = F_x(-\infty) = 0$

(3) $\lim_{b \to +\infty} F_x(b) = F_x(+\infty) = 1$

ใช้คำจำกัดความของเหตุการณ์ E_b เหตุการณ์ของรูป $\{a < X \leqslant b\}$ สามารถอธิบายได้เป็นเซทของผลลัพธ์ w ใน sample space ดังเช่นตัวแปรเชิงสุ่ม X มีค่ามากกว่า a แต่ไม่เกิน b ดังนั้น $P(a < X \leqslant b)$ สามารถเขียนได้เป็น

$$F_x(b) - F_x(a)$$

ในกรณีเกี่ยวกับเวลามาถึงของลูกค้าคนแรกของแต่ละคนสองวัน sample space ประกอบด้วยจุดทั้งหมด (x_1, x_2) อย่างเช่น $0 \leqslant X_1, X_2 \leqslant 8$ พิจารณาเหตุการณ์ทั้งหมดที่สัมพันธ์กับการทดลองนี้และสมมติว่าความน่าจะเป็นของเหตุการณ์นั้นสามารถหาได้ สมมติว่า X เป็นส่วนเฉลี่ยของเวลามาถึงทั้งสองครั้งเป็นตัวแปรเชิงสุ่ม และ E_b เป็นเซทของผลลัพธ์ w ใน sample space ที่ประกอบขึ้นเป็นเหตุการณ์ E_b ดังเช่น $X \leqslant b$ ดังนั้น $F_x(b) = P(E_b) = P(X \leqslant b)$ สมมติว่า b = 4 ชั่วโมง ค่าทั้งหมดของ X_1, X_2 คำนวณหาได้ $(X_1 + X_2)/2 \leqslant 4$ หรือ $X_1 + X_2 \leqslant 8$ พื้นที่แลเงาในรูปที่ 1 ดังนั้น $F_x(b)$ เป็นเพียงความน่าจะเป็นที่เกิดขึ้นตามลำดับของเหตุการณ์กำหนดได้โดยพื้นที่แลเงา $F_x(b)$ สามารถคำนวณได้ถ้าหากว่าทราบความน่าจะเป็นของเหตุการณ์นั้นใน sample space



รูปที่ 1 พื้นที่แลเงาใช้แทนเหตุการณ์ $E_b = \{X \leqslant 4\}$

สมมติการทดลองกระทำซ้ำ ๆ กัน n ครั้งและตัวแปรเชิงสุ่ม X เป็นเวลาแต่ละครั้งที่ได้สังเกต แสดงได้ $X_1, X_2 \ldots X_n$ ผลลัพธ์ของการทดลอง n ครั้งเหล่านี้ ลำดับผลลัพธ์เหล่านี้ให้ X(1) เป็นค่าสังเกตที่เล็กที่สุด X(2) รองลงมา, ...X(n) เป็นค่าที่ใหญ่ที่สุด พลอตฟังก์ชันขั้นบันไดเหล่านี้ $F_n(x)$

สำหรับ $X < X_{(1)}$ ให้ $F_n(n) = 0$

สำหรับ $X(1) \leqslant X < X(2)$, ให้ $F_n(x) = 1/n$

สำหรับ $X(2) \leqslant X < X(3)$, ให้ $F_n(x) = 2/n$

สำหรับ $X(n-1) \leqslant X < X(n)$, ให้ $F_n(x) = (n-1)/n$

สำหรับ $X \geqslant X_n$, ให้ $F_n(x) = n/n = 1$

$F_n(x)$ สามารถตีความได้เหมือนเศษส่วนของผลลัพธ์ของการทดลองที่น้อยกว่าหรือเท่ากับ x และเรียกว่าฟังก์ชันของการแจกแจงของตัวอย่าง

ปัญหาส่วนมากในทางปฏิบัติ มีอยู่สิ่งหนึ่งที่ไม่ได้เกี่ยวข้องกับเหตุการณ์ใน sample space แต่มุ่งมายังตัวแปรเชิงสุ่มและฟังก์ชันการแจกแจง ดังตัวอย่าง ตัวแปรเชิงสุ่ม X_1 เวลาของการมาถึงครั้งแรกของวันแรกเป็น exponential ในทำนองเดียวกัน เวลาของการมาถึงครั้งแรกของวันที่สอง X_2 แล้ว ฟังก์ชันของการแจกแจงของตัวแปรเชิงสุ่ม $\bar{X} = (x_1 + x_2)/2$ สามารถคำนวณหาได้ เงื่อนไขเหล่านี้ออกมาในรูปของ C.D.F ไม่ได้ออกมาในรูปของความน่าจะเป็นของเหตุการณ์ใน sample space

ความน่าจะเป็นแบบมีเงื่อนไขและเหตุการณ์ที่มีความอิสระกัน

เมื่อไรการทดลองประกอบด้วยอุปสงค์สำหรับผลิตภัณฑ์ระหว่างแต่ละเดือนของสองเดือน อุปสงค์ระหว่างเดือนแรกได้มาเมื่อสิ้นสุดเดือนแรก หรือเวลามาถึงของลูกค้าสองคนแรกของแต่ละวันของสองวันเท่าที่ได้สังเกตตรงต่อเวลา รายละเอียดนี้สามารถนำไปใช้ทำนายผลของการทดลองต่อไป! รายละเอียดเช่นนั้นไม่ต้องการสัมพันธ์กับเวลา ถ้าหากว่าสำรวจอุปสงค์สำหรับสองผลิตภัณฑ์ระหว่างเดือนหนึ่ง โดยทราบอุปสงค์ของผลิตภัณฑ์ชนิดหนึ่งอาจเป็นประโยชน์ในการประเมินอุปสงค์ของอีกชนิดหนึ่ง เพื่อที่จะใช้รายละเอียดนี้มาเป็นแนวความคิดของความน่าจะเป็นแบบมีเงื่อนไข เพื่อที่จะนิยามเหตุการณ์ที่เกิดขึ้นใน sample space

พิจารณาสองเหตุการณ์ใน sample space E_1 กับ E_2 ในเมื่อ E_1 ใช้แทนเหตุการณ์ที่เกิดขึ้นแล้วและ E_2 ใช้แทนเหตุการณ์ที่ซึ่งจะเกิดหรือไม่เกิด สมมติว่า $P(E_1) > 0$ ความน่าจะเป็นแบบมีเงื่อนไขของการเกิดขึ้นของเหตุการณ์ E_2 กำหนดว่าเหตุการณ์ E_1 เกิดขึ้นแล้ว $P(E_2/E_1)$ นิยามได้เป็น

$$P(E_2/E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

ในเมื่อ $(E_1 \cap E_2)$ ใช้แทนเหตุการณ์ที่ประกอบด้วยจุดทั้งหมด w ใน sample space ทั้ง E_1 และ E_2 ร่วมกัน ดังตัวอย่าง พิจารณาการทดลองซึ่งประกอบด้วยอุปสงค์สำหรับผลิตภัณฑ์หนึ่งแต่ละเดือนของสองเดือน สมมติว่า sample space ประกอบด้วยจุดทั้งหมด $w = (X_1, X_2)$ ในเมื่อ X_1 ใช้แทนอุปสงค์ระหว่างเดือนแรกและ X_2 ใช้แทนอุปสงค์ระหว่างเดือนที่สอง $X_1, X_2 = 0, 1, \ldots 99$ เป็นที่ทราบว่าอุปสงค์สำหรับเดือนแรกเป็น 10 ดังนั้นเหตุการณ์ E_1 ประกอบด้วยจุด (10, 0), (10, 1), (10, 2),...(10, 99) เกิดขึ้นแล้ว พิจารณาเหตุการณ์ E_2 ซึ่งใช้แทนอุปสงค์สำหรับผลิตภัณฑ์ในเดือนที่สองซึ่งไม่เกินหนึ่งหน่วย เหตุการณ์นี้ประกอบด้วยจุด (0, 0), (1, 0), (2, 0),... (10, 0),..., (99, 0), (0, 1), (1, 1), (2, 1)...(10, 1),...(99,1) เหตุการณ์ $(E_1 \cap E_2)$ ประกอบด้วยจุด (10, 0) และ (10, 1) ดังนั้น ความน่าจะเป็นของอุปสงค์ซึ่งไม่เกินหนึ่งหน่วยในเดือนที่สอง กำหนดว่าอุปสงค์ของ 10 หน่วย เกิดขึ้นระหว่างเดือนแรก $P(E_2/E_1)$ กำหนดได้

$$P(E_2/E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{P(w = (10, 0), w = (10, 1))}{(P(w = (10, 0), w = (10, 1), \ldots w = (10, 99))}$$

คำจำกัดความของความน่าจะเป็นแบบมีเงื่อนไขสามารถกำหนดออกมาเป็นความถี่โดยให้ n เป็นจำนวนครั้งของการทดลอง และ n_1 เป็นจำนวนครั้งของเหตุการณ์ E_1 เกิดขึ้น

n_{12} เป็นจำนวนครั้งที่เหตุการณ์ $E_1 \cap E_2$ เกิดขึ้นในการทดลอง n_1 ครั้ง n_{12}/n_1 เป็นสัดส่วนของครั้งที่เหตุการณ์ E_2 เกิดขึ้นเมื่อ E_1 เกิดขึ้นแล้วด้วย n_{12}/n_1 เป็นความถี่สัมพันธ์แบบมีเงื่อนไขของ E_2 กำหนดว่า E_1 เกิดขึ้นแล้ว ความถี่สัมพัทธ์ n_{12}/n_1 เท่ากับ $(n_{12}/n)/(n_1/n)$ การใช้ความถี่มาแสดงความน่าจะเป็นเมื่อ n มีค่ามาก n_{12}/n คือ $P(E_1 \cap E_2)$ โดยประมาณ n_1/nคือ $P(E_1)$ โดยประมาณ ดังนั้นความถี่สัมพัทธ์แบบมีเงื่อนไขของ E_2 กำหนด E_1 คือ $P(E_1 \cap E_2)/P(E_1)$ โดยประมาณ

ความน่าจะเป็นแบบมีเงื่อนไขมีคุณสมบัติดังนี้

(1) $0 \leqslant P(E_2/E_1) \leqslant 1$

(2) ถ้า E_2 เป็นเหตุการณ์หนึ่งซึ่งไม่สามารถเกิดขึ้นแล้ว $P(E_2/E_1) = 0$

(3) ถ้าเหตุการณ์ E_2 เป็น Sample space S แล้ว $P(S/E_1) = 1$

(4) ถ้า E_2 และ E_3 เป็นเหตุการณ์ที่ไม่ร่วมกันใน S แล้ว

$$P(E_2 + E_3/E_1) = P(E_2/E_1) + P(E_3/E_1)$$

ถ้า $P(E_2) > 0$ แล้วเราสามารถกำหนด $P(E_1/E_2) = P(E_1 \cap E_2)/P(E_2)$

ถ้าหากว่ารายละเอียดของการเกิดขึ้นของเหตุการณ์ E_1 ไม่ได้ให้รายละเอียดของการเกิดขึ้นหรือไม่เกิดขึ้นของเหตุการณ์ E_2 แล้ว $P(E_2/E_1) = P(E_2)$ หรือ $P(E_1/E_2) = P(E_1)$ ก็กล่าวได้ว่า E_1 กับ E_2 เป็นเหตุการณ์ที่อิสระกัน ถ้าหากว่า E_1 กับ E_2 มีความอิสระกัน $P(E_1) > 0$ แล้ว $P(E_2/E_1) = P(E_1 \cap E_2)/P(E_1) = P(E_2)$ ดังนั้น $P(E_1 \cap E_2) = P(E_1)P(E_2)$ ดังตัวอย่าง ถ้าอุปสงค์สำหรับผลิตภัณฑ์ระหว่างเดือนหนึ่งทราบว่าไม่มีผลต่ออุปสงค์ในเดือนต่อไปแล้ว เหตุการณ์ E_1 กับ E_2 กล่าวได้ว่ามีความอิสระกัน ในกรณีนี้

$$P(E_2/E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

$$= \frac{P(w = (10, \ ), w = (10, 1))}{P(w = (10, 0), w = (10, 1), \ldots, w = (10, 99))}$$

$$= \frac{P(E_1)P(E_2)}{P(E_1)} = P(E_2)$$

คำจำกัดความของความอิสระกันสามารถขยายออกไปยังเหตุการณ์ใด ๆ $E_1, E_2, \ldots, E_n$ เป็นเหตุการณ์อิสระถ้าทุก ๆ กลุ่มย่อยของเหตุการณ์เหล่านี้ $E_1^*, E_2^*, \ldots, E_k^*$

$P(E_1^* \cap E_2^* \cap \ldots \cap E_k^*) = P(E_1^*)\, P(E_2^*) \ldots P(E_k^*)$ ที่สอดคล้องกับการเกิดขึ้นของ

การแจกแจงความน่าจะเป็นแบบไม่ต่อเนื่อง

ตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องเป็นค่าของเซทที่จำกัดฟังก์ชันการแจกแจงสำหรับตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง $F_X(b)$ กำหนดได้โดย

$$F_X(b) = P(X(w) \leqslant b) = \sum_{\text{all } X_i \leqslant b} P(X(w) = X_i)$$

ในเมื่อ $(X(w) = X_i)$ เป็นเซทของผลลัพธ์ w ใน sample space อย่างเช่นตัวแปรเชิงสุ่ม X มีค่าเป็น X_i ดังตัวอย่าง ตัวแปรเชิงสุ่ม X ใช้แทนกำลังสองของอุปสงค์ให้ N = 99 ตัวแปรเชิงสุ่มอาจเป็นค่าหนึ่งค่าใดของเลข 0, 1, 4, 9, ..., $(98)^2$, $(99)^2$ ดังนั้น $X_0 = 0, X_1 = 1, X_2 = 4, X_3 = 9, \ldots,$ $X_{98} = (98)^2$ $X_{99} = (99)^2$ และ $\{X(w) = X_0\} = \{X(w) = 0\}, \{X(w) = X_1\} = \{X(w) = 1\}, \ldots$ $\{X(w) = X_{98}\} = \{X(w) = (98)^2\}, \{X(w) = X_{99}\} = \{X(w) = (99)^2\}$ เราเขียนเสียใหม่ได้ $\{X(w) = X_i\}$ สำหรับ X_i ทั้งหมด ดังนั้น

$$\begin{aligned} F_X(150) &= \sum_{\text{all } X_i \leqslant 150} P(X(w) = X_i) \\ &= P(X(w) = X_0) + P(X(w) = X_1) + \ldots + P(X(w) = X_{12}) \\ &= P(X(w) = 0) + P(X(w) = 1 + \ldots + P(X(w) = 144) \\ &= \sum_{\text{all } k \leqslant 150} P(X(w) = k) \end{aligned}$$

สำหรับ $k = 0, 1, 4, 9, \ldots (98)^2, (99)^2$

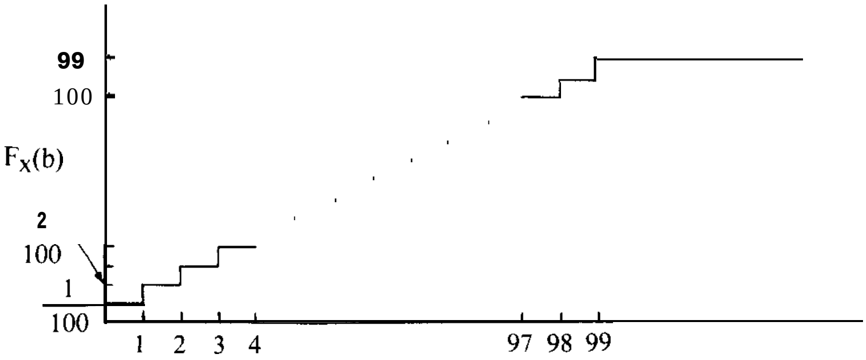
โดยทั่ว ๆ ไป

$$F_X(b) = \sum_{\text{all } X_i \leqslant (b)} P(X(w) = X_i) = \sum_{\text{all } k \leqslant (b)} P(X(w) = k)$$

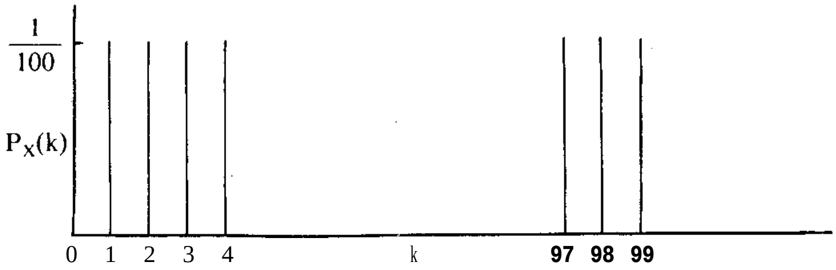
$P(X(w) = k)$ เขียนได้เป็น $P_X(k)$ ดังนั้น

$$F_X(b) = \sum_{\text{all } k \leqslant b} P_X(k)$$

$P_X(k)$ เรียกว่าการแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่มไม่ต่อเนื่อง X ดังตัวอย่างของเราตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องใช้แทนอุปสงค์สำหรับผลิตภัณฑ์ในเดือนหนึ่งที่กำหนดให้ N = 99 $P_X(k) = P(X = k) = 1/100$ สำหรับ $k = 0, 1, \ldots 99$, ทั้งหมดแล้ว C.D.F. สำหรับตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องนี้กำหนดให้ดังในรูปที่ 2 การแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องแสดงในรูปที่ 3



รูปที่ 2 C.D.F. ของตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง



รูปที่ 3 การแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง

แน่ละความสูงของเส้นแนวตั้งในรูปที่ 3 เท่ากันหมด เพราะว่า $P_X(0) = P_X(1) = P_X(2) = \ldots = P_X(99)$ สำหรับตัวแปรเชิงสุ่ม X อื่น ๆ $P_X(k)$ ไม่ต้องการเท่ากัน ดังนั้นเส้นแนวตั้งจะไม่คงที่ ผลบวกของการแจกแจงความน่าจะเป็น $P_X(k)$ เท่ากับหนึ่ง

$$\sum_{\text{all } b} P_X(k) = 1$$

มีหลาย ๆ การแจกแจงความน่าจะเป็นแบบไม่ต่อเนื่องที่ใช้กับงานการวิจัยการดำเนินงาน ซึ่งจะกล่าวในหัวข้อต่อไป

การแจกแจงทวินาม

ตัวแปรเชิงสุ่ม X มีการแจกแจงทวินามถ้าการแจกแจงความน่าจะเป็นสามารถเขียนได้เป็น

$$P(X = k) = P_X(k) = \frac{n!}{k!(n-k)!} P^k (1-P)^{n-k}$$

ในเมื่อ P เป็นค่าคงที่มีค่าอยู่ระหว่างศูนย์กับหนึ่ง n เป็นเลขจำนวนเต็มบวก และ k เป็นเลขจำนวนเต็มบวกด้วย $0 \leqslant k \leqslant n$ สังเกตว่า $P_X(k)$ มีค่าเป็นบวกเสมอและพิสูจน์ได้ว่า

$$\sum_{k=0}^{n} P_X(k) = 1$$

การแจกแจงของฟังก์ชันนี้ประกอบด้วยสองพารามิเตอร์ n และ P การแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่มนี้แสดงในรูปที่ 4 เมื่อไร n = 1 การแจกแจงทวินามในกรณีนี้

$$P(X = 0) = P_X(0) = 1 - P$$

และ $$P(X = 1) = P_X(1) = P$$

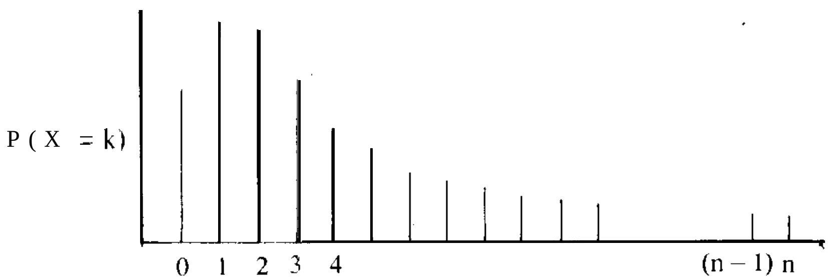
เรียกตัวแปรเชิงสุ่มนี้ว่าการแจกแจงแบบเบอร์นูลลี่ ดังนั้นถ้าหากว่าตัวแปรเชิงสุ่มหนึ่งมีสองค่า 0 กับ 1 ด้วยความน่าจะเป็น $1 - P$ กับ P ตามลำดับ

ถ้าหากว่า $X_1, X_2, \ldots, X_n$ เป็นตัวแปรเชิงสุ่มแบบเบอร์นูลลีที่มีความอิสระกัน แต่ละตัวแปรมีพารามิเตอร์ P แล้วเราสามารถแสดงตัวแปรเชิงสุ่ม

$$X = X_1 + X_2 + \ldots + X_n$$

เป็นตัวแปรเชิงสุ่มแบบทวินามมีพารามิเตอร์ n กับ P ดังนั้น ถ้าเราโยนเหรียญที่สมดุลหนึ่งอัน 10 ครั้ง จำนวนหัวทั้งหมด $X = X_1 + X_2 + \ldots + X_{10}$ มีการแจกแจงแบบทวินามพร้อมด้วยพารามิเตอร์ 10 กับ 1/2

$$P(X = k) = \frac{10!}{k!(10-k)!} (1/2)^k (1/2)^{10-k}$$



รูปที่ 4 การแจกแจงทวินามเมื่อกำหนดให้ n และ P คงที่

การแจกแจงแบบพัวซอง

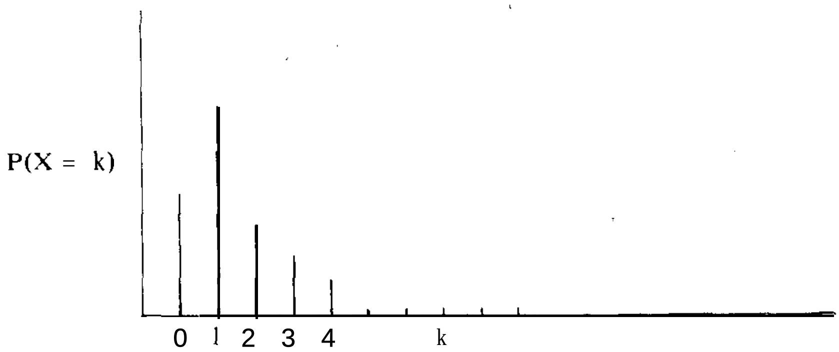
ตัวแปรเชิงสุ่ม X มีการแจกแจงแบบพัวซองถ้าหากว่าการแจกแจงความน่าจะเป็นสามารถเขียนได้เป็น

$$P(X = k) \quad = \quad P_X(k) \;=\; \frac{\lambda^k e^{-\lambda}}{k!}$$

ในเมื่อ λ เป็นค่าคงที่บวก (เป็นพารามิเตอร์ของการแจกแจงนี้ 1 และ k เป็นเลขจำนวนเต็มบวก $P_X(k)$ มีค่าเป็นบวกและแสดงได้ว่า

$$\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \quad = \quad 1$$

ตัวอย่างการแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่มแบบพัวซองดังแสดงในรูปที่ 5



รูปที่ 5 การแจกแจงแบบพัวซอง

ในการวิจัยการดำเนินงานจะต้องใช้การแจกแจงแบบพัวซองบ่อย ๆ การแจกแจงนี้เหมาะสมหลาย ๆ สภาวะที่เหตุการณ์เกิดขึ้นตลอดระยะเวลาคล้ายกับการมาถึงของลูกค้า เมื่อไรที่เหตุการณ์คล้ายกับเหตุการณ์ที่จะเกิดขึ้นในช่วงหนึ่งเหมือนในช่วงอื่น ๆ ของเหตุการณ์เกิดขึ้นจะไม่มีผลไม่ว่าเหตุการณ์อื่น ๆ จะเกิดขึ้นหรือไม่ ดังนั้น จำนวนของลูกค้ามาถึงในเวลาที่กำหนดเป็นการแจกแจงแบบพัวซอง ในทำนองเดียวกัน อุปสงค์สำหรับผลิตภัณฑ์กำหนดให้ก็สมมติให้เป็นการแจกแจงนี้

การแจกแจงแบบเรขาคณิต

ตัวแปรเชิงสุ่ม X มีการแจกแจงแบบเรขาคณิตถ้าหากว่าการแจกแจงความน่าจะเป็นสามารถเขียนได้เป็น

$$P(X = k) = P_X(k) = p(1 - p)^{k-1}$$

ในเมื่อพารามิเตอร์ p เป็นค่าคงที่มีค่าระหว่าง 0 กับ 1 และ k มีค่า 1, 2, 3,... $P_X(k)$ มีค่าเป็นบวกและ

$$\sum_{k=1}^{\infty} p(1-p)^{k-1} = 1$$

การแจกแจงแบบเรขาคณิตที่มีประโยชน์ในสภาวะดังนี้ สมมติการดำเนินการทดลองหนึ่งนำไปสู่การลำดับของตัวแปรเชิงสุ่มแบบเบอร์นูลลีที่อิสระกันซึ่งแต่ละตัวแปรมีพารามิเตอร์ p อย่างเช่น $P(X_i = 1) = p$ และ $P(X = 0) = 1 - p$ สำหรับ i ทั้งหมด ตัวแปรเชิงสุ่ม X (เป็นจำนวนของการทดลองที่เกิดขึ้นจนกระทั่งตัวแปรเชิงสุ่มแบบเบอร์นูลลีแรกมีค่าเป็น 1) มีการแจกแจงแบบเรขาคณิตด้วยพารามิเตอร์ p

การแจกแจงความน่าจะเป็นแบบต่อเนื่อง

ตัวแปรเชิงสุ่มแบบต่อเนื่องได้ถูกนิยามเป็นตัวแปรเชิงสุ่มซึ่งมีค่าต่อเนื่องกัน C.D.F สำหรับตัวแปรเชิงสุ่ม แบบต่อเนื่อง $F_X(b)$ สามารถเขียนได้เป็น

$$F_X(b) = P(X(w) \leqslant b) = \int_{\infty}^{b} f_X(y)dy$$

ในเมื่อ $f_X(y)$ เป็นตัวฟังก์ชันของความหนาแน่นของตัวแปรเชิงสุ่ม X subscript X ใช้แสดงตัวแปรเชิงสุ่ม เมื่อไรที่ไม่มีความยุ่งยาก subscript อาจตัดทิ้งได้ $f_X(y)$ จะเขียนได้เป็น $f(y)$ ฟังก์ชันความหนาแน่นสามารถใช้คำนวณความน่าจะเป็นทุกชนิด

$$P(a < X \leqslant b) = F(b) - F(a) = \int_{a}^{b} f_X(y)dy$$

$P(a < X \leqslant b)$ มีความสัมพันธ์กับความน่าจะเป็นที่ผลลัพธ์ w ของการทดลองใน sample space อย่างเช่นเมื่อไรก็ตามเหตุการณ์ที่ $X(w)$ อยู่ระหว่าง a กับ b โดยทั่ว ๆ ไปฟังก์ชันความหนาแน่นสามารถหาได้จาก C.D.F โดยใช้ความสัมพันธ์

$$\frac{dF_X(y)}{dy} = \frac{d}{dy}\int_{-\infty}^{y} f_X(t) \; dt = f_X(y)$$

ฟังก์ชันความหนาแน่นกำหนดได้ในรูปที่ 6 พื้นที่ภายใต้ฟังก์ชันความหนาแน่นทั้งหมดมีค่าเป็นหนึ่งเสมอ

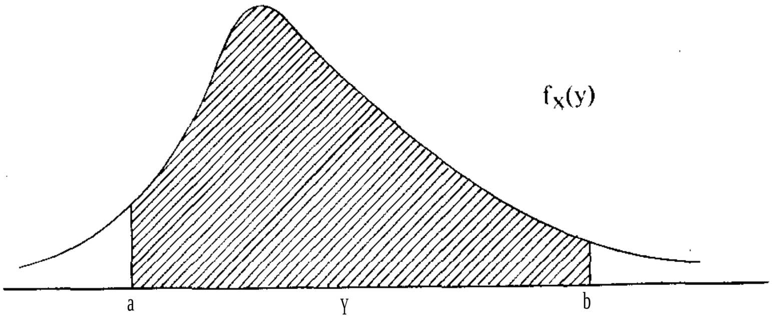
พื้นที่ภายใต้ฟังก์ชันระหว่าง a กับ b เป็นเพียง

$$P(a \leqslant X \leqslant b)$$

เมื่อ a = b พื้นที่ภายใต้เส้นโค้งเป็นศูนย์ ดังนั้น $P(a \leqslant X \leqslant b) = P(X = a) = 0$ เพราะฉะนั้น ตัวแปรเชิงสุ่มแบบต่อเนื่อง a และ b ใด ๆ

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

แต่นี้จะไม่เป็นความจริงสำหรับตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง



รูปที่ 6 ฟังก์ชันความหนาแน่นของตัวแปรเชิงสุ่ม

ในการนิยาม C.D.F สำหรับตัวแปรเชิงสุ่มแบบต่อเนื่องก็เช่นเดียวกันกับ $f_X(y)$ ที่นิยามให้เป็นค่าของ y จาก $-\infty$ ถึง $+\infty$

$$F_X(b) = \int_{-\infty}^{b} f_X(y)\, ay$$

นี้ไม่เป็นสาเหตุของความยากลำบาก ถึงแม้ว่าตัวแปรเชิงสุ่มไม่สามารถมีค่าลบหรือถูกจำกัดต่อพื้นที่อื่น ๆ (อย่างเช่น เวลามาถึงของลูกค้าคนแรก) เนื่องจากว่า $f_X(y)$ สามารถเป็นศูนย์เหนือส่วนของช่วงจาก $-\infty$ ถึง $+\infty$ ข้อบังคับของฟังก์ชันความหนาแน่นมีอยู่ 2 ข้อ

(1) $f_X(y)$ มีค่าเป็นลบไม่ได้

(2) $\int_{-\infty}^{\infty} f_X(y)\,dy = 1$

ดังนั้น $f_X(y)$ ไม่สามารถให้เป็น $P(X = y)$ ได้ เนื่องจากว่าความน่าจะเป็นนี้เป็นศูนย์ อย่างไรก็ตาม $f_X(y)\,dy$ เป็นความน่าจะเป็นที่ตัวแปรเชิงสุ่ม X วางอยู่ในช่วง $(y, y+dy)$ $f_X(y)$ เป็นการวัดความถี่ซึ่งตัวแปรเชิงสุ่มจะตกอยู่ภายในช่วงเล็ก ๆ ใกล้กับ y

มีอยู่หลาย ๆ การแจกแจงความน่าจะเป็นแบบต่อเนื่องที่สำคัญใช้ในงานวิจัยการดำเนินดังจะได้กล่าวต่อไป

การแจกแจงแบบเอกซ์โพเนนเชียล

ตัวแปรเชิงสุ่มแบบต่อเนื่องหนึ่ง ซึ่งมีความหนาแน่นกำหนดได้

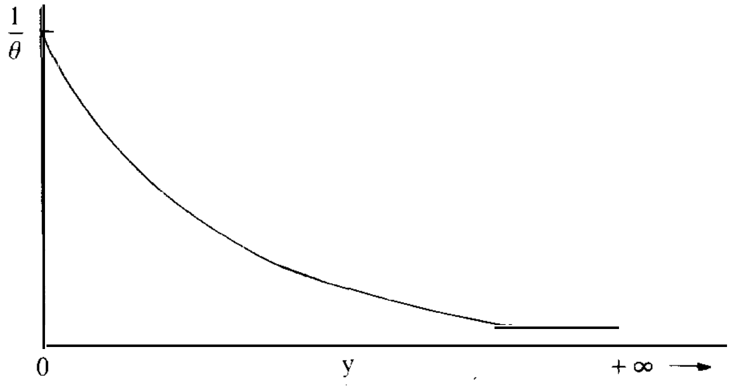
$$f_X(y) = \begin{cases} \frac{1}{\theta}e^{-y/\theta} & \text{สำหรับ } y \geq 0 \\ 0 & \text{สำหรับ } y < 0 \end{cases}$$

เป็นที่ทราบกันว่าตัวแปรเชิงสุ่มที่มีการแจกแจงแบบเอกซ์โพเนนเชียลเป็นฟังก์ชันหนึ่งที่มีพารา-มิเตอร์เดียว θ ในเมื่อ θ เป็นค่าคงที่ $f_X(y)$ เป็นฟังก์ชันความหนาแน่น

เนื่องจากว่ามีค่าไม่เป็นลบ integrate all range มีค่าเท่ากับหนึ่ง

$$\int_{-\infty}^{\infty} f_X(y)dy = \int_0^{\infty} \frac{1}{\theta}e^{-y/\theta}\,dy = -e^{-y/\theta}\Big|_0^{\infty} = 1$$

ฟังก์ชันความหนาแน่นแบบ exponential ดังแสดงได้ในรูปที่ 7



รูปที่ 7 ฟังก์ชันความหนาแน่นแบบเอกซ์โพเนนเชียล

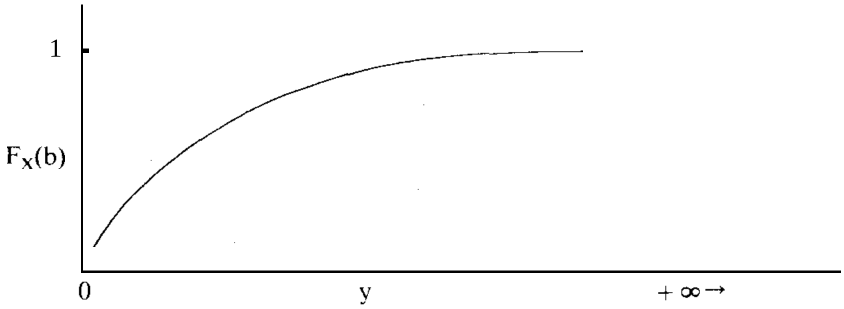
C.D.F ของตัวแปรเชิงสุ่มของการแจกแจงแบบเอกซ์โพเนนเชียล $F_X(b)$ กำหนดได้โดย

$$F_X(b) = \int_{-\infty}^{b} f_X(y)\,dy$$

$$= \begin{cases} 0 & \text{สำหรับ } b < 0 \\ \int_0^b \frac{1}{\theta}e^{-y/\theta}dy = 1 - e^{-\theta/b} & \text{สำหรับ } b \geq 0 \end{cases}$$

และแสดงในรูปที่ 8

การแจกแจงแบบเอกซ์โพเนนเชียลใช้กันอย่างกว้างขวางในการวิจัยการดำเนินงาน เวลาระหว่างลูกค้ามาถึง ระยะเวลาของการสนทนาทางโทรศัพท์ อายุของส่วนประกอบหลอดอิเล็กทรอนิกส์



รูปที่ 8 C.D.F ของการแจกแจงแบบเอกซ์โพเนนเชียล

ดังตัวอย่าง สมมติว่าอายุของหลอดสุญญากาศสมมติให้มีการแจกแจงแบบเอกซ์โพเนนเชียล ถ้าหากว่าหลอดมีอายุมานาน 1000 ชั่วโมง ความน่าจะเป็นของการมีอายุมานานเพิ่มอีก 50 ชั่วโมงเหมือนกับความน่าจะเป็นของการมีอายุเพิ่มอีก 50 ชั่วโมง กำหนดว่าหลอดมีอายุมานาน 2000 ชั่วโมง หรือจะกล่าวได้อีกแบบหนึ่งว่า หลอดสุญญากาศยี่ห้อใหม่ไม่ดีกว่าหลอดที่ใช้มาแล้ว 1000 ชั่วโมง ความหมายของเอกซ์โพเนนเชียลนั้นมีความสำคัญมากและพบอยู่เสมอในทางปฏิบัติ

การแจกแจงแบบแกมม่า

ตัวแปรเชิงสุ่มแบบต่อเนื่องซึ่งฟังก์ชันความหนาแน่นเขียนได้

$$f_a(y) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}} y^{(\alpha-1)}e^{-y/\beta} & \text{สำหรับ } y \geq 0 \\ 0 & \text{สำหรับ } y < 0 \end{cases}$$

เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบแกมม่า ความหนาแน่นนี้เป็นฟังก์ชันสองพารามิเตอร์ α และ β ทั้งสองตัวเป็นค่าคงที่บวกและ $\Gamma(\alpha)$ นิยามได้เป็น

$$\Gamma(\alpha) = \int_0^{\infty} t^{\alpha-1}e^{-t}\,dt \qquad \text{สำหรับ} \quad \alpha > 0\text{ทั้งหมด}$$

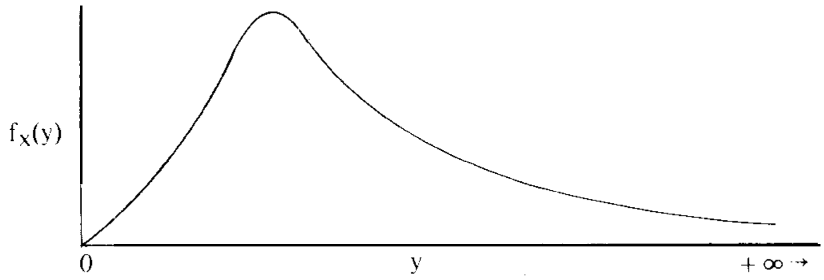
ถ้าหากว่า α เป็นเลขจำนวนเต็มบวกแล้ว integrate by part ซ้ำ ๆ กันจะได้

$$\Gamma(\alpha) = (\alpha - 1)! = (\alpha - 1)(\alpha - 2) \quad \ldots\ldots\ldots 3.2.1$$

กราฟของฟังก์ชันความหนาแน่นแบบแกมม่ากำหนดได้ในรูปที่ 9 ตัวแปรเชิงสุ่มที่มีความหนาแน่นแบบแกมม่ามีประโยชน์ใช้แทนคณิตศาสตร์ของปรากฏการณ์ทางกายภาพ อย่างเช่นเวลาของการบริการลูกค้ามีการแจกแจงแบบเอกซ์โพเนนเชียลด้วยพารามิเตอร์ θ ตัวแปรเชิงสุ่ม T เป็นเวลาทั้งหมดที่ใช้บริการลูกค้า n คน แล้วจะมีการแจกแจงแบบแกมม่าด้วยพารามิเตอร์ n กับ θ (ใช้แทน α กับ β ตามลำดับ)

$$P(T < t) = \int_0^t \frac{1}{\Gamma(n)\theta^n} y^{n-1} e^{-y/\theta} dy$$

เมื่อไรที่ n = 1 (หรือ $\alpha = 1$) ฟังก์ชันความหนาแน่นแบบแกมม่าจะเป็นฟังก์ชันความหนาแน่นของตัวแปรเชิงสุ่มแบบเอกซ์โพเนนเชียล ดังนั้น ผลบวกของตัวแปรเชิงสุ่มที่มีการแจกแจงแบบเอกซ์โพเนนเชียลที่อิสระกันจะมีการแจกแจงแบบแกมม่า



รูปที่ 9 ฟังก์ชันความหนาแน่นแบบแกมม่า

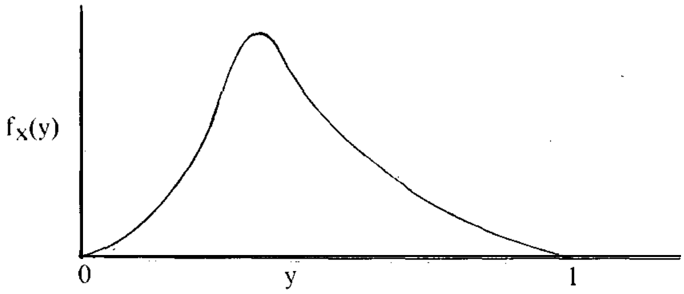
การแจกแจงที่สำคัญอีกแบบหนึ่งคือการแจกแจงแบบไค-สแคว ซึ่งมีความสัมพันธ์กับการแจกแจงแบบแกมม่าถ้าหากว่า X เป็นตัวแปรเชิงสุ่มมีการแจกแจงแบบแกมม่าพร้อมด้วยพารามิเตอร์ $\beta = 1$ และ $\alpha = \nu/2$ (ν เป็นเลขจำนวนเต็มบวก) แล้วตัวแปรเชิงสุ่มใหม่ Z = 2X มีการแจกแจงแบบไค-สแคว พร้อมด้วยองศาแห่งความอิสระ

การแจกแจงแบบเบต้า

ตัวแปรเชิงสุ่มแบบต่อเนื่องซึ่งฟังก์ชันความหนาแน่นเขียนได้

$$f_X(y) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{(\alpha-1)}(1-y)^{(\beta-1)} & \text{สำหรับ } 0 \leq y \leq 1 \\ 0 & \text{สำหรับค่าอื่น ๆ} \end{cases}$$

เป็นตัวแปรเชิงสุ่มทีมีการแจกแจงแบบเบต้า ความหนาแน่นที่เป็นฟังก์ชันของสองพารามิเตอร์ α กับ β ทั้งสองตัวเป็นค่าคงที่บวก กราฟของฟังก์ชันความหนาแน่นแบบเบต้ากำหนดได้ในรูปที่ 10



รูปที่ 10 ฟังก์ชันความหนาแน่นแบบเบต้า

รูปการแจกแจงแบบเบต้ามีประโยชน์เมื่อตัวแปรเชิงสุ่มถูกจำกัดในช่วงหนึ่งหน่วย (unit interval) โดยเฉพาะเมื่อไรที่ $\alpha = \beta = 1$ เรียกการแจกแจงแบบเบต้าว่าการแจกแจงแบบยูนิฟอร์มตลอดช่วงหน่วย ฟังก์ชันความหนาแน่นแสดงในรูปที่ 11 และมีโอกาสเกิดขึ้นระหว่าง 0 กับ 1 เท่ากันหมด

C.D.F. สำหรับตัวแปรเชิงสุ่มกำหนดได้โดย

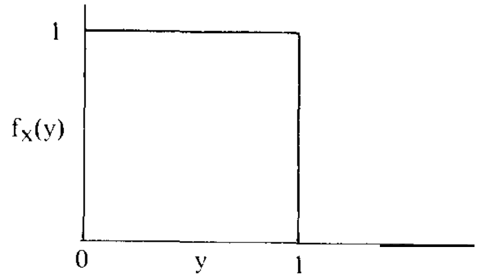
$$F_X(b) = \begin{cases} 0 & \text{สำหรับ } b < 0 \\ b & \text{สำหรับ } 0 \leqslant b \leqslant 1 \\ 1 & \text{สำหรับ } b > 1 \end{cases}$$

ถ้าหากว่าฟังก์ชันความหนาแน่นเป็นค่าคงที่ตลอดช่วง ๆ หนึ่งอย่างเช่น $[c, d]$ การแจกแจงแบบยูนิฟอร์มตลอดช่วงนี้สามารถหาได้ ฟังก์ชันความหนาแน่นกำหนดได้โดย

$$f_X(y) = \begin{cases} \dfrac{1}{d-c} & \text{สำหรับ } c \leqslant y \leqslant d \\ 0 & \text{สำหรับค่าอื่น ๆ} \end{cases}$$

ตัวแปรเชิงสุ่มนี้กล่าวได้ว่าเป็นการแจกแจงแบบยูนิฟอร์มตลอดช่วง $[c, d]$

การแจกแจงแบบ t ก็เป็นการแจกแจงที่สำคัญและมีความสัมพันธ์กับการแจกแจงแบบเบต้า ถ้าหากว่า X เป็นตัวแปรเชิงสุ่มมีการแจกแจงแบบเบต้าพร้อมด้วยพารามิเตอร์ $\alpha = 1/2$ และ $\beta = \nu/2$ (ν เป็นเลขจำนวนเต็มบวก) แล้วตัวแปรเชิงสุ่มใหม่ $Z = \sqrt{\nu X(1-X)}$ มีการแจกแจงแบบ t พร้อมด้วยองศาแห่งความอิสระ V



รูปที่ 11 การแจกแจงแบบยูนิฟอร์มตลอดช่วงหน่วย

การแจกแจงแบบปกติ

การแจกแจงหนึ่งที่สำคัญที่สุดในการวิจัยการดำเนินงานคือการแจกแจงแบบปกติ ตัวแปรเชิงสุ่มแบบต่อเนื่องที่มีฟังก์ชันความหนาแน่นกำหนดได้โดย

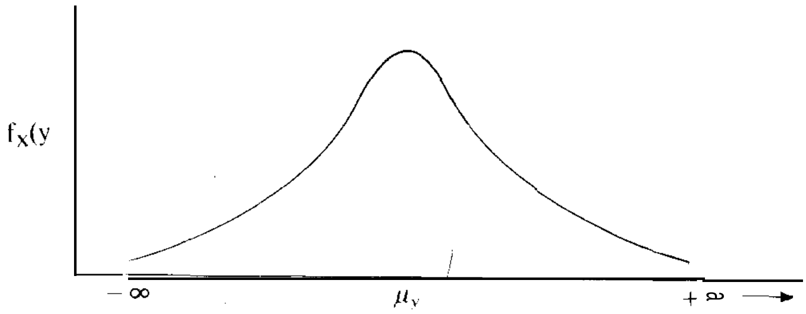
$$f_X(y) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \quad \text{สำหรับ } -\infty < y < \infty$$

เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบปกติ ฟังก์ชันความหนาแน่นมีสองพารามิเตอร์ μ กับ σ ในเมื่อ μ เป็นตัวคงที่และ σ มีค่าเป็นบวกกราฟของฟังก์ชันความหนาแน่นแบบปกติกำหนดได้ในรูปที่ 12 ฟังก์ชันความหนาแน่นนี้เป็นเส้นโค้งรูประฆังนั่นคือสมมาตรรอบ μ C.D.F. ของตัวแปรเชิงสุ่มที่มีการแจกแจงแบบปกติกำหนดได้โดย

$$F_X(b) = \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy$$

ทำการแปลง $Z = (Y-\mu)/\sigma$ C.D.F. สามารถเขียนได้

$$F_X(b) = \int_{-\infty}^{(b-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} dz$$



รูปที่ 12 ฟังก์ชันความหนาแน่นแบบปกติ

ฟังก์ชันนี้ไม่สามารถ integrate ได้ แต่ค่าที่คำนวณได้ในตารางคำนวณมาจากฟังก์ชันของ

$$\alpha = \int_{K_\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} dz$$

เป็นฟังก์ชันของ K_α ดังนั้น การหา $F_X(b)$ ในตารางโดยให้ $K_\alpha = (b-\mu)/\sigma$ และ

$$\alpha = \int_{K_\alpha}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} dz$$

$F_X(b)$ คำนวณได้จาก $1-\alpha$ ดังนั้น ถ้า $P(14 < X \leqslant 18) = F_X(18) - F_X(14)$ ในเมื่อ X มีการแจกแจงแบบปกติพร้อมด้วย $\mu = 10$ และ $\sigma = 4$ จากตารางเราได้ $(18-10)/4 = 2$ และคำนวณ $1-F_X(18) = 0.0228$ ในทำนองเดียวกันจากตาราง $(14-10)/4 = 1$ และคำนวณ $1-F_X(14) = 0.1587$ จากตัวเลขเหล่านี้หาค่า $F_X(18)-F_X(14) = 0.1359$ ถ้าหากว่า K_α มีค่าเป็นลบเราสามารถใช้ความสมมาตรของการแจกแจงปกติเนื่องจากว่า

$$F_X(b) = \int_{-\infty}^{(b-\mu)/\sigma} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} dz = \int_{-(b-\mu)/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-Z^2/2} dz$$

ในกรณีนี้ $-(b-\mu)/\sigma$ มีค่าเป็นบวกและ $F_X(b) = \alpha$ อ่านได้จากตารางโดยใช้ $-(b-\mu)/\sigma$ อย่างเช่นเราต้องการคำนวณหา

$$P(2 < X \leqslant 18) = F_X(18) - F_X(2)$$

$F_X(18)$ มีค่าเท่ากับ $1-0.0228 = 0.9772$ การคำนวณ $F_X(2)$ สังเกตว่า $(2-10)/4 = -2$ มีค่าเป็นลบ ดังนั้น จากตารางเราได้ $K_\alpha = +2$ และ $F_X(2) = 0.0228$ ดังนั้น

$$F_X(18) - F_X(2) = 0.9772 - 0.0228 = 0.9544$$

การแจกแจงแบบปกติสามารถแสดงว่า ถ้า $X_1, X_2, \ldots X_n$ เป็นตัวแปรเชิงสุ่มแบบปกติที่มีความอิสระกันพร้อมด้วยพารามิเตอร์ $(\mu_1, \sigma_1), (\mu_2, \sigma_2), \ldots (\mu_n, \sigma_n)$ ตามลำดับแล้ว $X = X_1 + X_2 + \ldots + X_n$ ก็เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบปกติด้วยพารามิเตอร์

$$\sum_{i=1}^{n} \mu_i$$

และ

$$\sqrt{\sum_{i=1}^{n} \sigma_i^2}$$

ถ้าหากว่า $X_1, X_2, \dots X_n$ ไม่มีการแจกแจงแบบปกติแล้ว ภายใต้ weak Conditions

$$X = \sum_{i=1}^{n} X_i$$

มีแนวโน้มเข้าสู่การแจกแจงแบบปกติขณะ n มีค่ามาก

ถ้า C เป็นค่าคงที่ใด ๆ และ X เป็นปกติด้วยพารามิเตอร์ μ และ σ แล้วตัวแปรเชิงสุ่ม CX มีการแจกแจงแบบปกติพร้อมด้วยพารามิเตอร์ $C\mu$ กับ $C\sigma$ ดังนั้น ถ้า $X_1, X_2, \dots X_n$ เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบปกติที่มีความอิสระกันแต่ละตัวแปรมีพารามิเตอร์ μ กับ σ แล้วตัวแปรเชิงสุ่ม $\bar{X} = \sum_{i=1}^{n} X_i/n$ จะมีการแจกแจงแบบปกติพร้อมด้วยพารามิเตอร์ μ กับ $\sigma/\sqrt{n}$

ค่าคาดหวัง

มีอยู่ค่าหนึ่งซึ่งอาจแสดงออกของตัวแปรเชิงสุ่มและการแจกแจงความน่าจะเป็นของตัวแปรที่เราต้องการ ค่านั้นก็คือค่าที่คาดหวังของตัวแปรเชิงสุ่ม อย่างเช่นใครคนหนึ่งพูดขึ้นว่าค่าที่คาดหวังของอุปสงค์สำหรับผลิตภัณฑ์ ค่าที่คาดหวังของเวลาที่ลูกค้าคนแรกมาถึงเป็นต้น

ค่าที่คาดหวังของตัวแปรเชิงสุ่ม X แสดงได้ E(X) และกำหนดได้โดย

$$E(X) = \begin{cases} \sum_{\text{all } k} k\, P(X = k) & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง} \\ \int_{-\infty}^{\infty} y\, f_X(y)\, dy & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบต่อเนื่อง} \end{cases}$$

สำหรับตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องจะเห็นได้ว่า E(X) เป็นเพียงผลบวกของผลิตภัณฑ์ของค่าที่อาจเป็นไปได้ที่ตัวแปรเชิงสุ่ม X มีค่าและสมนัยกับความน่าจะเป็น ดังตัวอย่างอุปสงค์ของผลิตภัณฑ์ที่ $k = 0, 1, 2, \dots 98, 99$ และ $P_X(k)$ สำหรับ i ทั้งหมด ค่าที่คาดหวังของอุปสงค์คือ

$$E(X) = \sum_{k=0}^{99} k\, P_X(k) = \sum_{k=0}^{99} k \frac{1}{100} = 49.5$$

ถ้าหากว่า X เป็นตัวแปรเชิงสุ่มแบบทวินามพร้อมด้วยพารามิเตอร์ n กับ p ค่าที่คาดหวังของ X กำหนดได้โดย

$$E(X) = \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k(1-p)^{n-k}$$

$$= np$$

ถ้าตัวแปรเชิงสุ่ม X มีการแจกแจงแบบพัวซองพร้อมด้วยพารามิเตอร์ λ E(X) คำนวณหาได้โดย

$$E(X) = \sum_{k=0}^{\infty} k \frac{\lambda^k e^{-\lambda}}{k!}$$

$$= \lambda$$

ถ้าตัวแปรเชิงสุ่ม X มีการแจกแจงแบบเรขาคณิตพร้อมด้วยพารามิเตอร์ p

$$E(X) = \sum_{k=1}^{\infty} k\, p(1-p)^{k-1}$$

$$= 1/p$$

สำหรับตัวแปรเชิงสุ่มแบบต่อเนื่อง ค่าที่คาดหวังก็สามารถคำนวณหาได้ ถ้า X มีการแจกแจงแบบเอกซ์โพเนนเชียลพร้อมด้วยพารามิเตอร์ θ ค่าที่คาดหวังกำหนดได้โดย

$$E(X) = \int_{-\infty}^{\infty} y f_X(y)dy = \int_0^{\infty} y \frac{1}{\theta} e^{-y/\theta} dy$$

$$= \theta$$

ถ้าตัวแปรเชิงสุ่ม X มีการแจกแจงแบบแกมม่าพร้อมด้วยพารามิเตอร์ α และ β ค่าที่คาดหวังของ X กำหนดได้โดย

$$\int_{-\infty}^{\infty} y f_X(y)dy = \int_0^{\infty} y \frac{1}{\Gamma(\alpha)\beta^{\alpha}} y^{(\alpha-1)} e^{-y/\beta} dy$$

$$= \alpha\beta$$

ถ้าหากว่าตัวแปรเชิงสุ่ม X มีการแจกแจงแบบเบต้าพร้อมด้วยพารามิเตอร์ α กับ β ค่าที่คาดหวังของ X กำหนดได้โดย

$$\int_{-\infty}^{\infty} y f_X(y)dy = \int_0^{1} y \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{(\alpha-1)}(1-y)^{(\beta-1)} dy$$

$$= \frac{\alpha}{\alpha+\beta}$$

สุดท้าย ถ้าตัวแปรเชิงสุ่ม X เป็นการแจกแจงแบบปกติพร้อมด้วยพารามิเตอร์ μ กับ σ ค่าที่คาดหวังของ X กำหนดได้โดย

$$\int_{-\infty}^{\infty} y\, f_X(y)\,dy = \int_{-\infty}^{\infty} y \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2} dy = \mu$$

ความคาดหวังของตัวแปรเชิงสุ่มมีประโยชน์มาก มีความหมายในรูปของส่วนเฉลี่ยของตัวอย่าง โดยเฉพาะถ้าตัวแปรเชิงสุ่มได้มาโดยการสังเกตซ้ำ ๆ กันและคำนวณมัชฌิมเลขคณิต $\bar{X}$ $\bar{X}$ จะเข้าสู่ความคาดหวังของตัวแปรเชิงสุ่ม X ขณะที่จำนวนของการทดลองมีค่ามาก

ความคาดหวังไม่ได้จำกัดแก่ตัวแปรเชิงสุ่ม X เท่านั้น ถ้าหากว่า Z เป็นบางฟังก์ชันของ X สมมติให้เป็น $Z = g(X)$ แล้ว $g(X)$ ก็เป็นตัวแปรเชิงสุ่มด้วย ความคาดหวังของ $g(X)$ สามารถนิยามได้เป็น

$$E[g(x)] = \begin{cases} \sum_{\text{all } k} g(k)\, P(X = k) & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง} \\ \int_{-\infty}^{\infty} g(y) f_X(y)\,dy & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบต่อเนื่อง} \end{cases}$$

ดังนั้น ความคาดหวังของ Z สามารถคำนวณได้โดยการใช้คำจำกัดความในรูปของความหนาแน่นของ Z หรือในทางกลับกันโดยการใช้คำจำกัดความที่เป็นความคาดหวังของฟังก์ชัน X เทียบกับฟังก์ชันความหนาแน่นของ X

โมเมนต์

ถ้าหากว่าฟังก์ชัน g ที่ได้อธิบายในหัวข้อก่อน ๆ กำหนดได้โดย

$$Z = g(X) = X^j$$

ในเมื่อ j เป็นเลขจำนวนเต็มบวกแล้ว ความคาดหวังของ X^j เรียกว่าโมเมนต์ที่ j รอบจุดกำเนิดของตัวแปรเชิงสุ่ม X และกำหนดได้โดย

$$E(X^j) = \begin{cases} \sum_{\text{all } K} K^j P_X(k) & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง} \\ \int_{-\infty}^{\infty} y^j f_X(y)\,dy & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบต่อเนื่อง} \end{cases}$$

เมื่อไร่ $j = 1$ โมเมนต์ที่หนึ่งคือความคาดหวังของ X แสดงด้วยสัญลักษณ์ μ เรียกว่ามัชฌิมเลขคณิตของการแจกแจง

การใช้ทฤษฎีของ Unconscious Statistician เพื่อคำนวณความคาดหวังของ $Z = g(X) = CX$ ในเมื่อ C เป็นค่าคงที่ ถ้า X เป็นตัวแปรเชิงสุ่มแบบต่อเนื่องแล้ว

$$E|CX| = \int_{-\infty}^{\infty} Cyf_X(y)dy = C\int_{-\infty}^{\infty} yf_X(y)dy = \mathbf{CE(X)}$$

ดังนั้น ความคาดหวังของผลคูณของตัวคงที่กับตัวแปรเชิงสุ่มจะมีค่าเท่ากับตัวคงที่คูณด้วยความคาดหวังของตัวแปรเชิงสุ่มนี้เป็นจริงสำหรับตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องด้วย

ถ้าหากว่าฟังก์ชัน g ที่ได้อธิบายในหัวข้อก่อน ๆ กำหนดได้โดย $Z = g(X) = (X-E(X))^j = (X-\mu)^j$ ในเมื่อ j เป็นเลขจำนวนเต็มบวกแล้ว ความคาดหวังของ $(X-\mu)^j$ เรียกว่าโมเมนต์ที่ j รอบมัชฌิมเลขคณิตของตัวแปรเชิงสุ่ม X และกำหนดได้โดย

$$E(X-\mu)^j = \begin{cases} \sum_{\text{all } k} (k-\mu)^j P_X(k) & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง} \\ \int_{-\infty}^{\infty} (y-\mu)^j f_X(y)dy & \text{ถ้า X เป็นตัวแปรเชิงสุ่มแบบต่อเนื่อง} \end{cases}$$

สังเกตว่าถ้า $j = 1$ แล้ว $E(X-\mu) = 0$ ถ้า $j = 2$ แล้ว $E(X-\mu)^2$ เรียกว่าความแปรปรวนของตัวแปรเชิงสุ่ม X และแสดงได้ด้วย 2 รากของความแปรปรวน σ เรียกว่าส่วนเบี่ยงเบนมาตรฐานของตัวแปรเชิงสุ่ม X ในรูปของคำจำกัดความ $2 = E(X-\mu)^2 = E(X^2)-\mu^2$

ดังที่ได้แสดงมาแล้วว่าถ้า $Z = g(X) = CX$ แล้ว $E|CX| = CE(X) = C\mu$ ในเมื่อ C เป็นค่าคงที่และ μ คือ E(X) ความแปรปรวนของตัวแปรเชิงสุ่ม $Z = g(X) = CX$ คำนวณหาได้ง่าย จากคำจำกัดความ ถ้า X เป็นตัวแปรเชิงสุ่มแบบต่อเนื่อง ความแปรปรวนของ Z กำหนดได้โดย $E(Z-E(Z))^2 = E(CX-CE(X))^2 = \int_{-\infty}^{\infty} (Cy-C\mu)^2 f_X(y)dy$

$$= C^2\int_{-\infty}^{\infty} (y-\mu)^2 f_x(y)dy = C^2\sigma^2$$

ดังนั้น ความแปรปรวนของผลคูณของตัวคงที่กับตัวแปรเชิงสุ่มมีค่าเท่ากับค่าคงที่กำลังสองคูณด้วยความแปรปรวนของตัวแปรเชิงสุ่ม นี้เป็นจริงสำหรับตัวแปรเชิงสุ่มแบบไม่

ต่อเนื่องด้วย ความแปรปรวนของตัวคงที่มีค่าเท่ากับศูนย์

มัชฌิมเลขคณิตและความแปรปรวนของตัวแปรเชิงสุ่มมีประโยชน์มากในการวิจัยการดำเนินงาน สังเกตว่าบางตัวแปรเชิงสุ่มมีมัชฌิมเลขคณิตให้ลักษณะแสดงออกที่สมบูรณ์ของการแจกแจงอย่างการแจกแจงแบบพัวซอง ส่วนการแจกแจงแบบปกติมีมัชฌิมเลขคณิตและความแปรปรวนให้ลักษณะแสดงออกที่สมบูรณ์ของการแจกแจง ส่วนการแจกแจงอื่น ๆ มีลักษณะบางอย่างแสดงออกเฉพาะตัวของการแจกแจง

การทำการตัดสินใจโดยปราศจากข้อมูล

เรามาพิจารณาปัญหาการเก็บรักษาสินค้า กองทัพอากาศต้องการสั่งซื้อเครื่องบินชนิดใหม่และจำนวนเครื่องอะไหล่ กองทัพอากาศต้องสั่งเครื่องอะไหล่เหล่านี้เป็นชุดละ 5 และสามารถเลือกระหว่างเครื่องอะไหล่ 15, 20 หรือ 25 บริษัทขายเครื่องเหล่านี้มีสองโรงงาน กองทัพอากาศต้องทำการตัดสินใจก่อนที่จะทราบว่าจะใช้ของโรงงานไหน จากประสบการณ์ในอดีตทราบว่าจำนวนเครื่องอะไหล่ที่ต้องการมีลักษณะเป็นการแจกแจงแบบพัวซองเมื่อไรที่การผลิตได้กระทำที่โรงงาน A ด้วยพารามิเตอร์ $\lambda = 21$ และโรงงาน B ด้วยพารามิเตอร์ $\lambda = 24$ มูลค่าเครื่องอะไหล่ที่ซื้อปัจจุบัน \$250,000 ถ้าซื้อในวันต่อมาจะเป็น \$500,000 โดยไม่คิดต้นทุนและดอกเบี้ย เครื่องอะไหล่ต้องมีสำรองอยู่เสมอ เมื่อต้องการและเครื่องที่ไม่ใช้ก็จะกลายเป็นเศษเหล็กเมื่อไรที่เครื่องบินพ้นสมัย

ก่อนการคำนวณหาคำไขต่อปัญหานี้ จะต้องสร้างสูตรโครงสร้างของงานทั่ว ๆ เพื่อทำการตัดสินใจ ผู้ทำการตัดสินใจต้องเลือกพฤติกรรมหนึ่ง a จากเซท A ที่เป็นไปได้ เซท A ประกอบด้วยสามจุดด้วยกัน a_1, a_2, a_3 ที่สมนัยกับการสั่งเครื่องอาหลั่ย 15, 20 หรือ 25 ตามลำดับ การที่จะเลือกพฤติกรรม ผู้ทำการตัดสินใจต้องทราบเหตุผลของพฤติกรรมซึ่งเป็นฟังก์ชันของสภาวะธรรมชาติ สภาวะธรรมชาติ θ ใช้แทนสภาวะที่เป็นไปได้ว่าจะใช้พฤติกรรมอันไหน โดยทั่ว ๆ ไปสภาวะธรรมชาติใช้แทนแบบจำลองความน่าจะเป็นของปรากฏการณ์ทางกายภาพที่กำลังศึกษาอยู่และแสดงออกด้วยพารามิเตอร์ของกลุ่มการแจกแจงความน่าจะเป็น เซทของมูลค่าที่ θ สมมติให้เป็น (H) ซึ่งประกอบด้วยสองจุด θ_1 กับ θ_2 ซึ่งสามารถเป็น $\lambda = 21$ กับ $\lambda = 24$ เพื่อที่จะวัดผลที่จะติดตามมาของพฤติกรรมของผู้ทำการตัดสินใจ จะต้องสมมติฟังก์ชันขาดทุน $l(a, \theta)$ ที่หาค่าได้ซึ่งสะท้อนการขาดทุนจากพฤติกรรม a เมื่อสภาวะธรรมชาติเป็น θ ถ้าปัญหาการสร้างสูตรในรูปผลกำไร ผลกำไรก็สามารถอยู่ในรูปของขาดทุนที่มีเครื่องหมาย

ลบโดยทั่ว ๆ ฟังก์ชันขาดทุนหาค่าออกมาในรูปของเงินถึงแม้ว่าสามารถใช้ฟังก์ชันยูทีลิตี้ อย่างเช่นถือเอาพฤติกรรม a_2 สภาวะธรรมชาติเป็น θ_1 ขึ้นอยู่กับจำนวนของเครื่องอะไหล่ที่ต้องการจริง ๆ จำนวนของเครื่องอะไหล่ที่ต้องการจริง ๆ เป็นตัวแปรเชิงสุ่มมีการแจกแจงความน่าจะเป็นเป็นฟังก์ชันของ $\lambda = 21$ ในสภาวะนี้ การวัดการขาดทุนในรูปของค่าที่คาดหวัง (expected cost) และเป็นฟังก์ชันของ a_2 กับ θ_1 เพราะฉะนั้นเมื่อไรตัวแปรเชิงสุ่มที่มูลค่าสมนัยกับ a และ θ การคำนวณหาค่าฟังก์ชันขาดทุนให้หาออกมาในรูปค่าที่คาดหวัง

ตารางที่ 1 ตารางขาดทุนสำหรับปัญหาการเก็บรักษาสินค้า

พฤติกรรม	สภาวะธรรมชาติ	
	$\theta_1 : \lambda = 21$	$\theta_2 : \lambda = 24$
a_1 : สั่ง 15	6.8265×10^6	8.265×10^6
a^* : สั่ง 20	6.178×10^6	7.270×10^6
a_3 : สั่ง 25	6.514×10^6	7.002×10^6

สำหรับตัวอย่างของเราตารางขาดทุนสามารถคำนวณได้ ค่าในตารางที่ถือเอาพฤติกรรม a_2 เมื่อสภาวะธรรมชาติเป็น θ_1 คำนวณได้ดังนี้ ให้ตัวแปรเชิงสุ่ม Z ใช้แทนมูลค่าของเครื่องอะไหล่ที่ซื้อเป็นฟังก์ชันของอุปสงค์ ดังนั้นเมื่อไรถือเอา a_2 และสภาวะธรรมชาติจริงเป็น $\theta_1 = 21$ $(\lambda = 21)$ แล้ว

$$Z(D) = \begin{cases} (250{,}000)(20) + (500{,}000)(D-20) & \text{สำหรับ } D > 20 \\ (250{,}000)(20) & \text{สำหรับ } D \leqslant 20 \end{cases}$$

ในเมื่อ D จำนวนเครื่องอะไหล่ที่ต้องการ เป็นตัวแปรเชิงสุ่มมีการแจกแจงแบบพัวซองพร้อมด้วยพารามิเตอร์ $\lambda = 21$

มูลค่าของเครื่องอะไหล่ที่ซื้อทั้งหมดที่คาดหวังในตารางขาดทุนใช้เนื้อหาของหัวข้อที่ได้ศึกษาข้างต้น มูลค่าที่คาดหวังของเครื่องอะไหล่ที่ซื้อ E(Z) สามารถคำนวณได้

$$E(Z) = \sum_{k=0}^{\infty} Z(k)P_D(k) = \sum_{k=0}^{20} Z(k)P_D(k) + \sum_{k=21}^{\infty} Z(k)P_D(k)$$

$$= \sum_{k=0}^{20} 5{,}000{,}000\, P_D(k) + \sum_{k=21}^{\infty} |5{,}000{,}000 + 500{,}000\, k - 10{,}000{,}000|\, P_D(k)$$

$$= 5{,}000{,}000\ P(D \leqslant 20) - 5{,}000{,}000\, P(D > 20) + 500{,}000 \sum_{k=21}^{\infty} k\, P_D(k)$$

$$= -5{,}000{,}000 + 10{,}000{,}000\, P(D \leqslant 20) + 500{,}000 \sum_{k=21}^{\infty} \frac{k(21)^k e^{-21}}{k!}$$

ในเมือ $P(D \leqslant 20)$ คือความน่าจะเป็นของตัวแปรเชิงสุ่มแบบพัวซองพร้อมด้วยพารามิเตอร์ $\lambda = 21$ ที่น้อยกว่าหรือเท่ากับ 20 (ดูตาราง)

$$P(D \leqslant 20) = 0.471$$

และ

$$\sum_{k=21}^{\infty} \frac{k(21)^k e^{-21}}{k!} = 21 \sum_{k=21}^{\infty} \frac{(21)^{(k-1)} e^{-21}}{(k-1)!}$$

$$= 21 \sum_{j=20}^{\infty} \frac{(21)^j e^{-21}}{j!}$$

$$= 21(1 - 0.384) = 12.936$$

ดังนั้น $E(Z) = -5{,}000{,}000 + 4{,}710{,}000 + 6{,}468{,}000 = 6{,}178{,}000$ ในทำนองเดียวกันมูลค่าอื่น ๆ อีกห้ามูลค่าในตารางก็คำนวณได้

จากทฤษฎียูทิลิตี้ regret function เป็นฟังก์ชันที่ตรงข้ามกับฟังก์ชันขาดทุน regret function $r(a, \theta)$ นิยามได้เป็น

$$r(a, \theta) = l(a, \theta) - \min_{a \in A} l(a, \theta)$$

การคำนวณ regret สำหรับ a และ θ ทั้งหมดได้โดยการลบด้วยขาดทุนที่น้อยที่สุดสำหรับ θ นี้จากขาดทุนสำหรับ a และ θ เฉพาะนี้ ตามธรรมดาฟังก์ชันขาดทุนกับฟังก์ชัน regret ไม่เท่ากัน

ตารางที่ 2 regret ตารางสำหรับปัญหาการเก็บรักษาสินค้า

พฤติกรรม	**สภาวะธรรมชาติ**	
	$\theta_1 : \lambda = 21$	$\theta_2 : \lambda = 24$
a_1 : สั่ง 15	$.6485 \times 10^6$	1.263×10^6
a_2 : สั่ง 20	0	$.268 \times 10^6$
a_3 : สั่ง 25	$.336 \times 10^6$	0

ปรัชญาการใช้ regret function แทนฟังก์ชันขาดทุนอธิบายได้ดังนี้ ถ้าหากว่าทราบสภาวะธรรมชาติจริงและผู้ทำการตัดสินใจได้พฤติกรรมที่ดีที่สุด ดังนั้น ตาราง regret จึงมีค่าหนึ่งเป็นศูนย์เสมอ สำหรับสภาวะธรรมชาติหนึ่งที่กำหนดให้ดูตารางที่ 2 สภาวะธรรมชาติจริง $\theta = 21$ พฤติกรรมที่ดีที่สุดเป็น a_2 สั่ง 20 รายการ ถ้าเลือกพฤติกรรม a_3 สั่ง 25 รายการ regret คือมูลค่า 0.336×10^6

บรรทัดฐานของเบยส์

ในบางสภาวะผู้ทำการตัดสินใจจะมีรายละเอียดบางอย่างล่วงหน้าเกียวกับ θ ทีขัดแย้งเกี่ยวกับเงื่อนไขนี้ว่าอะไรควรจะทำ เมื่อไรที่เกิดอย่างนี้ขึ้น ผู้ทำการตัดสินใจควรจะทำรายละเอียดนี้ลงในรายการ รายละเอียดเช่นนั้นสามารถแปรเป็นการแจกแจงความน่าจะเป็นทำหน้าที่เป็นสภาวะธรรมชาติการแจกแจงนี้เป็น prior distribution; prior distribution เป็น subjective ที่ขึ้นอยู่ประสบการณ์หรือเรียนรู้ของแต่ละชนิด

สำหรับปัญหาการเก็บรักษาสินค้าของเรา กองทัพอากาศอาจรู้จากประสบการณ์ในอดีตว่า 2/3 ของเครื่องอะไหล่ทุกชนิดผลิตในโรงงาน A และ 1/3 เท่านั้นที่ผลิตในโรงงาน B ดังนั้น prior distribution สำหรับ θ อาจสมมติได้เป็น

$$P(\theta = 21) = P_\theta(21) = \frac{2}{3}$$

$$P(\theta = 24) = P_\theta(24) = \frac{1}{3}$$

กระบวนการใช้ prior distribution ให้เป็นประโยชน์เพื่อเพิ่มการเลือกพฤติกรรมเป็นบรรทัดฐานของเบยส์ หลักของเบยส์บอกให้ผู้ทำการตัดสินใจเลือกพฤติกรรมนั้น (เรียกว่ากระบวนการตัดสินใจของเบยส์) ซึ่งทำให้ขาดทุนที่คาดหวังน้อยที่สุด การประเมินขาดทุนที่คาดหวังเทียบกับ prior distribution ซึ่งนิยามปกคลุมสภาวะธรรมชาติที่เป็นไปได้ ดังนั้น ขาดทุนที่คาดหวัง $E[l(a)]$ สำหรับแต่ละพฤติกรรมกำหนดได้โดย

$$E[l(a_1)] = (6.8265 \times 10^6)(2/3) + (8.265 \times 10^6)(1/3) = 7.306 \times 10^6$$

$$E[l(a_2)] = (6.178 \times 10^6)(2/3) + (7.270 \times 10^6)(1/3) = 6.542 \times 10^6$$

$$E[l(a_3)] = (6.514 \times 10^6)(2/3) + (7.002 \times 10^6)(1/3) = 6.677 \times 10^6$$

เพราะฉะนั้น หลักของเบยส์นำไปสู่การเลือกพฤติกรรม a_2 ผู้ทำการตัดสินใจควรจะปรับปรุงขาดทุนที่คาดหวังโดยใช้อุบายผสมมากกว่าอุบายบริสุทธิ์ แต่นี้ก็แสดงให้เห็นแล้วว่าผู้ทำการตัดสินใจไม่สามารถปรับปรุงสถานะของเขาโดยการใช้อุบายผสม ดังนั้น เป็นการเพียงพอสำหรับเขาที่จะพิจารณาอุบายบริสุทธิ์เท่านั้น

การทำการตัดสินใจโดยใช้ข้อมูล

ในหัวก่อนได้กล่าวถึงการตัดสินใจโดยปราศจากข้อมูลแต่ถ้าบางการทดลองที่เป็นไปได้ ข้อมูลได้มาจากการทดลองมีส่วนร่วมในกระบวนการตัดสินใจ ดังตัวอย่างในหัวข้อก่อนของเรา สมมติว่ารายละเอียดต่อไปนี้นำไปใช้ประโยชน์กับกองทัพอากาศ ชนิดของเครื่องก็เหมือนกันทุกอย่าง เครื่องอะไหล่ที่สั่งก็ผลิตจากโรงงานเดียวกันเหมือนก่อน ๆ แต่กองทัพอากาศไม่ทราบว่าสองโรงงานไหนที่ผลิต เหตุผลสำหรับขาดความรู้นี้เป็นเหตุให้จะต้องเร่งทำการตัดสินใจ กองทัพอากาศจึงต้องการข้อมูลเกี่ยวกับเครื่องอะไหล่จริง ๆ ในอดีต (เป็นการแจกแจงแบบพัวซอง) แต่ไม่มีเวลาที่จะกำหนดถิ่นฐานการผลิต ดูเหมือนว่ามีข้อดีสำหรับรายละเอียดที่เป็นประโยชน์นี้เข้าไปในกระบวนการตัดสินใจ ก่อนที่จะดำเนินตัวอย่าง จะขอกล่าววิธีการทั่ว ๆ ไปสำหรับข้อมูลที่รวมเข้าด้วยกัน

ให้ X เป็นรายละเอียดจากการทดลองได้มาจากตัวอย่างสุ่ม X เป็นตัวแปรเชิงสุ่มและอาจเป็นฟังก์ชันของข้อมูลตัวอย่าง X อาจเป็นมัชฌิมเลขคณิตของตัวอย่าง เวคเตอร์ของค่าสังเกตตัวอย่าง ค่าสังเกตตัวที่สามในตัวอย่าง ผู้ทำการตัดสินใจจะต้องเลือกฟังก์ชันตัดสินใจหรืออุบาย ซึ่งเป็นฟังก์ชันของ X ที่จะบอกผู้ทำการตัดสินใจว่าพฤติกรรมอะไร มีค่าแต่ละค่าที่เป็นไปได้ที่ X อาจเป็นฟังก์ชันนี้ แสดงได้เป็น $d(x)$ เพื่อว่าตัวแปรเชิงสุ่ม X มีค่า x แล้ว

$a = d(x)$ เป็นพฤติกรรมที่ได้ ผู้ทำการตัดสินใจสนใจเลือกฟังก์ชันที่ดีที่สุดจากหลาย ๆ ฟังก์ชันการตัดสินใจที่เป็นไปได้ การประเมินฟังก์ชันตัดสินใจจะต้องตรวจสอบผลที่จะติดตามมา เนื่องจากว่าพฤติกรรมที่ได้ a เป็นฟังก์ชันของผลลัพธ์ของตัวแปรเชิงสุ่ม X ขาดทุนมีความสัมพันธ์กับพฤติกรรมก็ขึ้นอยู่กับผลลัพธ์ของตัวแปรเชิงสุ่มนี้ การวัดผลที่เกิดมาภายหลังได้พฤติกรรม $a = d(x)$ เมื่อไรสภาวะธรรมชาติจริงเป็น θ กำหนดได้โดยขาดทุนที่คาดหวัง ปริมาณนี้ทราบได้ในลักษณะของฟังก์ชันการเสี่ยงภัย (risk function) $R(d, \theta)$

$$R(d, \theta) = |1(d, \theta)|$$

ในเมื่อความคาดหวังที่ได้เทียบกับการแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่ม X

พิจารณาการประยุกต์ในหัวข้อนี้ สมมติเราต้องการประเมินกฎการตัดสินใจ d_1 ถ้าต้องการจำนวนเครื่องอะไหล่จากข้อความในอดีตมากกว่า 23 สั่งเครื่องอะไหล่ 25 ชิ้น สำหรับเครื่องบินใหม่ สำหรับพฤติกรรมอื่น ๆ สั่ง 20 ชิ้น ดังนั้น ถ้าได้พฤติกรรม a_2 X (จำนวนเครื่องอะไหล่ที่ต้องการเกี่ยวกับข้อความในอดีต) คือ 23 หรือน้อยกว่า ถ้าได้พฤติกรรม a_3 X มีค่ามากกว่า 23 อย่างเช่น

$$a_2 = d_1(x) \quad \text{สำหรับ } X \leqslant 23$$

$$a_3 = d_1(x) \quad \text{สำหรับ } X > 23$$

เพราะฉะนั้น

$$\begin{aligned} R(d_1, 21) &= E|1(d_1, 21)| \\ &= (6.178 \times 10^6)P(X \leqslant 23) + (6.514 \times 10^6)P(X > 23) \\ &= 6.27 \times 10^6 \end{aligned}$$

ในเมื่อ X เป็นตัวแปรเชิงสุ่มแบบพัวซองพร้อมด้วยพารามิเตอร์ 21

$$\begin{aligned} R(d_1, 24) &= E|1(d_1, 24)| \\ &= (7.270 \times 10^6)P(X \leqslant 23) + (7.002 \times 10^6)P(X > 23) \\ &= 7.13 \times 10^6 \end{aligned}$$

ในเมื่อ X เป็นตัวแปรเชิงสุ่มแบบพัวซองพร้อมด้วยพารามิเตอร์ 24 ดังนั้นการประเมินฟังก์ชันการเสี่ยงภัยสำหรับฟังก์ชัน การตัดสินใจมีหลักการอย่างไร ฟังก์ชันการตัดสินใจที่ดีที่สุดถูกนิยามให้เป็นฟังก์ชันหนึ่งซึ่งให้การเสี่ยงภัยน้อยที่สุดสำหรับทุก ๆ ค่าของ θ เป็นที่แน่ใจว่าฟังก์ชันการตัดสินใจที่ดีที่สุด อาจหาค่าไม่ได้เสมอไป หรือส่วนมากหาค่าไม่ได้ คำว่าการเสี่ยงภัยใช้แสดงค่าที่คาดหวังของ regret

$$R(d, \theta) = E[r(d, \theta)]$$

มีหลาย ๆ กรณี ฟังก์ชันขาดทุนแสดงเป็น regret function ดังนั้น ฟังก์ชันการเสี่ยงภัยที่สมนัยกันก็เหมือนกัน ฟังก์ชันการเสี่ยงภัยแสดงค่าออกมาในรูปของค่าที่คาดหวังของ regret ดังตัวอย่างของเรา

$$R(d_1, 21) = E[r(d_1, 21)]$$
$$= (0)P(X \leqslant 23)+(.336 \times 10^6)P(X > 23) = 0.0954 \times 10^6$$

และ

$$R(d_1, 24) = E[r(d_1, 24)]$$
$$= (.268 \times 10^6)P(X \leqslant 23) + (0)P(X > 23) = 0.127 \times 10^6$$

กระบวนการของเบย์ส

เมื่อไรใช้ข้อมูลให้เป็นประโยชน์ได้ และไม่มีวิธีการที่ดีที่สุดในการเลือกกระบวนที่ดีที่สุดด้วยข้อมูลจึงจำเป็นจะต้องใช้ทฤษฎีของเกม ซึ่งเป็นข้อเสียเมื่อไรที่ไม่มีข้อมูลใช้ให้เป็นประโยชน์ แต่ผู้ทำการตัดสินใจมีรายละเอียดบางอย่างเกี่ยวกับสภาวะธรรมชาติซึ่งอธิบายได้ในรูปของ proir distribution แล้วหลักของเบย์สก็สามารถใช้กับฟังก์ชันการเสี่ยงภัย ถ้าหากว่าสภาวะธรรมชาติเป็นแบบไม่ต่อเนื่อง การเสี่ยงภัยของเบย์สที่สมนัยกับ prior probability distribution ของ θ $P_\theta(k)$ กำหนดได้

$$B(d) = \sum_{\text{all } k} R(d_1 k)P_v(k)$$

ถ้าหากว่าสภาวะธรรมชาติเป็นแบบต่อเนื่อง การเสี่ยงภัยของเบย์สที่สมนัยกับ prior probability density function ของ θ, $P_\theta(y)$ กำหนดได้โดย

$$B(d) = \int_{-\infty}^{\infty} R(d, y)\, P_\theta(y)\, dy$$

หลักของเบย์สบอกให้ผู้ทำการตัดสินใจเลือกฟังก์ชันนั้นซึ่งทำให้ B(d) มีค่าน้อยที่สุด วิธีการคำนวณหากระบวนการตัดสินใจของเบย์สดังเสนอต่อไปนี้

เมื่อไรที่ไม่ใช้ข้อมูลให้เป็นประโยชน์ กระบวนการของเบย์สได้เลือกพฤติกรรมที่ทำให้ขาดทุนที่คาดหวังน้อยที่สุด ความคาดหวังนี้ประเมินได้ด้วย prior distribution ของ θ แต่กรณีนี้นำข้อมูลมาใช้คือต้องการเครื่องอะไหล่ 30 รายการ นี้เป็นหลักฐานเกี่ยวกับเครื่องอะไหล่

ว่าได้ผลิตในโรงงาน A หรือ B ภายหลังได้สังเกตข้อมูลเหล่านี้ prior distribution ควรจะถูกปรับปรุงให้ทันสมัยเพื่อใช้รายละเอียดให้เหมาะยิ่งขึ้นเกี่ยวกับการแจกแจงความน่าจะเป็นของสภาวะธรรมชาติ รายละเอียดที่ปรับปรุงให้ทันสมัยนี้เรียก posterior distribution ของ θ กำหนด prior distribution และข้อมูล $X = x$ posterior distribution ของ θ เป็นการแจกแจงแบบมีเงื่อนไขของ θ กำหนด $X = x$ ถ้า θ เป็นแบบไม่ต่อเนื่อง posterior distribution แสดงได้ $h\theta/X = x(k)$ ถ้า θ เป็นแบบต่อเนื่อง posterior distribution แสดงได้ $h\theta/X = x(y)$ ในกรณีที่ใช้การแจกแจงความน่าจะเป็นแบบตัวแปรทวิและการแจกแจงความน่าจะเป็นแบบ marginal และเงื่อนไข (θ, X) เป็นตัวแปรเชิงสุ่มทวิมีการแจกแจงความน่าจะเป็นร่วม พิจารณากรณีที่ (θ, X) เป็นตัวแปรเชิงสุ่มทวิแบบไม่ต่อเนื่องมีการแจกแจงความน่าจะเป็นร่วมกำหนดได้ $P_{\theta\chi}(kj)$ ตัวแปรเชิงสุ่ม θ กับ X แต่ละตัวมีการแจกแจงแบบ marginal $P\theta(k)$ (prior distribution ของ θ) เป็นการแจกแจงแบบ marginal ของ θ โดยทั่ว ๆ ไปแสดงออกมาในรูปการแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่ม X ที่สมนัยกับการแจกแจงแบบมีเงื่อนไขของ X กำหนด θ จริง ๆ ดังตัวอย่าง ถ้า X มีการแจกแจงแบบพัวซองพร้อมด้วยพารามิเตอร์ $\lambda = 24$ แล้ว $e^{-24}24^j/j$ เป็นการแจกแจงความน่าจะเป็นแบบมีเงื่อนไขของ X

กำหนด $\lambda = 24$ ดังแสดงได้

$$Qx/\theta = k(j) = P(X = j/S = k)$$

ดังนั้น

$$Qx/\theta = 24(j) = P(X = j/S = 24) = \frac{e^{-24}24^j}{j!}$$

ถ้าหากว่าการแจกแจงร่วมของ (θ, X) เขียนได้เป็น

$$P\theta x(k, j) = Qx/\theta = k(j)P\theta(k)$$

และสามารถใช้คำนวณ $Qx(j)$ การแจกแจงแบบ marginal ของ X คำนวณได้

$$Qx(j) = \sum_{\text{all } k} P\theta x(kj) = \sum_{\text{all } k} Qx/\theta = k(j)P\theta(k)$$

การแจกแจงแบบ posterior ของ θ กำหนด $X = x$, $h\theta/X = x(k)$
(การแจกแจงแบบมีเงื่อนไขของ θ กำหนด $X = x$) และการแสดงออกของการแจกแจงความน่าจะเป็นส่วนของ (θ, X) ในทางกลับกันกำหนดได้โดย

$$P\theta x(k, j) = h\theta/X = x(k)\,Qx(j)$$

ซึ่งจะนำไปสู่การคำนวณการแจกแจงแบบ posterior

$$h\theta/X \;=\; x(k) \;=\; \frac{Qx/\theta = k(j)P\theta(k)}{Qx(j)}$$

ดังนั้นการคำนวณหาการแจกแจงแบบ posterior โดยใช้สมการข้างต้น $P\theta(k)$ เป็นการแจกแจงแบบ prior $Qx/\theta = k(j)$ เป็นการแสดงธรรมดาสำหรับการแจกแจงความน่าจะเป็นของตัวแปรเชิงสุ่ม X ฟังก์ชัน $Qx(j)$ เป็นการแจกแจงแบบ marginal ของตัวแปรเชิงสุ่ม X และคำนวณหาได้จาก

$$Qx(j) \;=\; \sum_{\text{all } k} Qx/\theta \;=\; k(j)P\theta(k)$$

กลับมาดูตัวอย่างของเรา สมมติว่าต้องการเครื่องอะไหล่ 30 รายการสำหรับครั้งก่อน ๆ prior distribution ของ θ คือ

$$P(\theta = 21) \;=\; P\theta(21) \;=\; \frac{2}{3}$$

$$P(\theta = 24) \;=\; P\theta(24) \;=\; \frac{1}{3}$$

การคำนวณหา $h\theta/x = 30(21)$ กับ $h\theta/x = 30(24)$ ในเมื่อ

$$h\theta/x \;=\; 30(21) \;=\; \frac{Qx/\theta = 21P\theta(21)}{Qx(30)}$$

และ

$$h\theta/x \;=\; 30(24) \;=\; \frac{Qx/\theta = 24(30)P\theta(24)}{Qx(30)}$$

ในกรณีนี้

$$Qx/\theta \;=\; 21(30) \;=\; P(X = 30/\theta = 21)$$

และ

$$Qx/\theta \;=\; 24(30) \;=\; P(X = 30/\theta = 24)$$

เป็นตัวแปรเชิงสุ่มแบบพัวซองมีค่า 30 เมื่อพารามิเตอร์เป็น 21 กับ 24 ตามลำดับ ความน่าจะเป็นเหล่านี้คำนวณหาได้จากตารางของการแจกแจงแบบพัวซอง

$$P(X = 30) \;=\; \sum_{k=0}^{30} Px(k) - \sum_{k=0}^{29} Px(k)$$

ดังนั้น

$$Qx/\theta \;=\; 21(30) \;=\; 0.013$$

และ

$$Qx/\theta = 24(30) = 0.036$$

ดังนั้น

$$Qx(30) = Qx/\theta = 21(30)P\theta(21) + Qx/\theta = 24(30)\ Pe(24)$$

$$= 0(.013)(2/3) + (0.036)(1/3) = 0.02067$$

เพราะฉะนั้น

$$h\theta/x = 30(21) = \frac{(0.013)(2/3)}{0.02067} = 0.420$$

$$h\theta/x = 30(24) = \frac{(0.036)(1/3)}{0.02067} = 0.580$$

จะเห็นได้ว่า ผลของการสังเกตบางข้อมูล (อุปสงค์ครั้งก่อน ๆ สำหรับเครื่องอะไหล่ 30 รายการ) prior probability ที่ใช้โรงงาน (θ = 21) ผลิตเครื่องอะไหล่ลดลงจาก 2/3 เป็น posterior probability 0.420 ในทำนองเดียวกันใช้ posterior probability ที่โรงงาน B(θ = 24) เพื่อผลิตเครื่องอะไหล่เพิ่มขึ้นจาก 1/3 เป็น posterior probability 0.580 กระบวนการของเบย์สเพื่อหาขาดทุนที่คาดหวังเมื่อเทียบกับ posterior distribution ของ θ สำหรับแต่ละพฤติกรรมดังนี้

$$E|1(a_1)| = (.6485\times10^6)(.42) + (1.263\times10^6)(.58) = 1.005\times10^6$$

$$E|1(a_2)| = (0)(.42) + (.268\times10^6)(.58) = .155\times10^6$$

$$E|1(a_3)| = (.336\times10^6)(.42) + (0)(.58) = .141\times10^6$$

กระบวนการของเบย์สควรเลือกพฤติกรรม a_3 (เนื่องจากว่าทำให้ขาดทุนที่คาดหวังน้อยที่สุด) ซึ่งสั่งเครื่องอะไหล่ 25 รายการ จะเห็นได้ว่าข้อมูลจากการทดลองจะเปลี่ยนพฤติกรรมของผู้ทำการตัดสินใจ

เท่าที่ได้แสดงมา posterior distribution ที่กำหนดให้นั้น θ และ X เป็นตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง ถ้าหากว่า θ เป็นแบบไม่ต่อเนื่องและ X เป็นแบบต่อเนื่อง posterior distribution $h\theta/X = x(x)$ กำหนดได้

$$he/X = x(k) = \frac{f_X/\theta = k(x)P\theta(k)}{f_X(x)}$$

ในเมื่อ $P\theta(k)$ เป็น prior distribution ของ θ และ $f_X/\theta\, k^{(x)}$ เป็นฟังก์ชันความหนาแน่นของตัวแปรเชิงสุ่ม X แต่เขียนในรูปนี้เพื่อแสดงว่าขึ้นอยู่กับค่าของพารามิเตอร์ θ ฟังก์ชัน $f_X(x)$ เป็น marginal density ของตัวแปรเชิงสุ่ม X และคำนวณหาได้จาก

$$f_X(x) = \sum_{\text{all } k} f_X/\theta(k)^{(x)}\ PO'“'$$

ถ้าหากว่าทั้ง θ กับ X เป็นแบบต่อเนื่อง posterior distribution $h\theta/X = x(y)$ กำหนดได้โดย

$$h\theta/X = x(y) = \frac{f_X/\theta = y(x)\ P\theta(y)}{f_X(x)}$$

ในเมื่อ $P\theta(y)$ เป็น prior density function ของ θ, นิยามฟังก์ชัน $f_X/\theta = y^{(x)}$ เหมือนหัวข้อก่อน ๆ $f_X(x)$ เป็น marginal density ของตัวแปรเชิงสุ่ม X และคำนวณหาได้จาก

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

ถ้าหากว่า θ เป็นแบบต่อเนื่องและ X เป็นแบบไม่ต่อเนื่อง posterior distribution $h\theta/X = x(y)$ กำหนดได้โดย

$$h\theta/X = x^{(y)} = \frac{Qx/\theta = y(j)\ P\theta(y)}{Qx(j)}$$

ในเมื่อ $P\theta(y)$ เป็น prior density function ของ θ ฟังก์ชัน $Qx/\theta = y(j)$ เป็นการแจกแจงความน่าจะเป็นตัวแปรเชิงสุ่ม X อย่างเช่น $P(X = j / \theta = y)$ แต่เขียนได้ในรูปนี้ขึ้นอยู่กับค่าของพารามิเตอร์ θ $Qx(j)$ เป็นการแจกแจงแบบ marginal ของตัวแปรเชิงสุ่ม X และคำนวณได้จาก

$$Qx(j) = \int_{-\infty}^{\infty} Qx/\theta = y(j)\, P\theta(y)\, dy$$

แบบฝึกหัด

1. อายุ X (ชั่วโมง) ของหลอดวิทยุมีฟังก์ชันความหนาแน่นน่าจะเป็นกำหนดได้ดังนี้

$$f_X(y) = \begin{cases} \dfrac{100}{y^2} & \text{สำหรับ } y \geqslant 100 \\ 0 & \text{สำหรับ } y < 100 \end{cases}$$

(ก) จงคำนวณหาความน่าจะเป็นที่หลอดวิทยุจะมีอายุ 150 ชั่วโมง

(ข) จงคำนวณค่าที่คาดหวังของตัวแปรเชิงสุ่ม

2. ให้ X เป็นตัวแปรเชิงสุ่ม มีฟังก์ชันความหนาแน่น

$$f_X(y) = \begin{cases} K(1-y^2) & \text{สำหรับ } -1 < y < 1 \\ 0 & \text{สำหรับค่าอื่น ๆ} \end{cases}$$

(ก) จงคำนวณหาค่า K ที่จะทำให้ $f_X(y)$ เป็นฟังก์ชันความหนาแน่นจริง ๆ

(ข) จงหาค่า C.D.F. ของ X

(ค) จงหา $E(2X-1)$

(ง) จงหาค่าแปรปรวนของ X

(จ) จงคำนวณหาค่าของ $P(\bar{X} > .05)$ โดยประมาณในเมื่อ $\bar{X}$ เป็นมัชฌิมเลขคณิตของตัวอย่าง จากตัวอย่างสุ่มขนาด $n = 100$ จากการแจกแจงข้างต้น (n มีค่าใหญ่)

3. อายุของทรานซิสเตอร์ X (เป็นชั่วโมง) มีการแจกแจงดัง

$$f_X(y) = \begin{cases} \dfrac{\alpha}{10{,}000}\left(1-\dfrac{y}{10{,}000}\right) & 0 < y < 10{,}000 \\ 0 & \text{ค่าอื่น ๆ} \end{cases}$$

(ก) จงคำนวณหาค่า α

(ข) จงคำนวณหาค่าที่คาดหวังของอายุทรานซิสเตอร์

(ค) จงคำนวณ C.D.F. สำหรับฟังก์ชันความหนาแน่นนี้

(ง) ถ้าหากว่า X ใช้แทนตัวแปรเชิงสุ่มของอายุทรานซิสเตอร์ ให้ $z = 2X$ ตัวแปรเชิงสุ่มใหม่ใช้ผลของ (a) หา C.D.F. ของ z

4. ในกระบวนการเคมีหนึ่ง มีขวดอยู่สามขวดที่ใช้บรรจุของเหลวมาตรฐานซึ่งบรรจุอยู่ในภาชนะใหญ่ การศึกษาแต่ละขวดแสดงมัชฌิมเลขคณิตของการบรรจุ 6 ออนซ์ ส่วนเบี่ยงเบนมาตรฐาน 0.07 ออนซ์ ถ้าหากว่าทั้งสามขวดอยู่ในลักษณะตัวอย่างสุ่ม

(ก) จงคำนวณหาค่าที่คาดหวังและส่วนเบี่ยงเบนมาตรฐานของปริมาตรของเหลวในภาชนะใหญ่

(ข) ถ้าหากว่าการบรรจุแต่ละขวดเป็นการแจกแจงปกติความน่าจะเป็นที่ปริมาตรของเหลวในภาชนะใหญ่จะมากกว่า 48.25 ออนซ์เป็นเท่าไร

5. ทราบว่าการแจกแจงของอายุของหลอดไฟเป็นปกติมีมัชฌิมเลขคณิต μ และส่วนเบี่ยงเบนมาตรฐาน 200 ชั่วโมง มูลค่าทั้งหมดของหลอดไฟ 1,000 หลอด เป็น (1,000)(1/5,000)μ บาท ตัวอย่างสุ่มหนึ่งประกอบด้วย n หลอดที่ผู้ซื้อเลือกคิดเป็นเงิน 1,000(1/5,000) $\bar{X}$ บาท จงคำนวณหาขนาดของ n เพื่อว่าตามค่าจะเป็น 0.95 ที่ผู้ซื้อจะต้องไม่ขายเกินหรือขาดไปกว่า 20 บาท

6. สมมติว่ามีเหรียญที่ไม่สมดุลอยู่ 2 เหรียญ เหรียญที่ 1 มีความน่าจะเป็นที่ปรากฏหัว 0.3 และเหรียญที่ 2 มีความน่าจะเป็นที่จะปรากฏหัว 0.6 ความน่าจะเป็นที่จะโยนเหรียญที่ 1 เป็น 0.6 เหรียญที่ 2 เป็น 0.4 matrix ขาดทุนกำหนดได้

	โยนเหรียญที่ 1	**โยนเหรียญที่ 2**
a_1 : โยนเหรียญที่ 1	0	1
a_2 : โยนเหรียญที่ 2	1	0

(ก) กระบวนการของเบย์ส (พฤติกรรม) ก่อนการโยนเหรียญจะเป็นเท่าไร

(ข) กระบวนการของเบย์สถ้าโยนเหรียญหนึ่งครั้งและผลลัพธ์เป็นหัวจะเป็นเท่าไร และถ้าผลลัพธ์เป็นก้อยกระบวนการของเบย์สจะเป็นเท่าไร

7. มีฟิล์มใหม่ชนิดหนึ่งสำหรับกล้องถ่ายรูป 35 มม. ห่อด้วยแผ่นกระดาษ 5 แผ่น แต่ละแผ่นให้ตัวอย่างของรูปถ่าย เนื่องจากว่าเป็นกระบวนการใหม่ ผู้ผลิตได้ลงมือเพิ่มกระดาษห่อเพื่อทดสอบการเก็บรักษาที่จะจำหน่าย ในการส่งเสริมฟิล์มมีข้อเสนอที่จะคืนทุนทั้งหมดถ้าหากว่าหนึ่งของห้าแผ่นที่ห่อเกิดเสีย ราคาขายกำหนดที่ $ 1.00 ถ้าการรับประกันนี้สมบูรณ์ร้านถ่ายรูปจะขายฟิล์มราคา 50 Cents ถ้าการประกันข้างต้นจะชดใช้ 10 Cents ต่อแผ่นที่เสียหนึ่งแผ่น ต้นทุนของฟิล์มของร้านถ่ายรูป 25 Cents และไม่มีการคืน ร้านถ่ายรูปอาจเลือกได้ 3 พฤติกรรม

a_1 : ทิ้งฟิล์มเสีย

a_2 : ขายฟิล์ม \$ 1.00

a_3 : ขายฟิล์ม 50 Cents

(ก) ถ้าสภาวะธรรมชาติทั้งหกสมนัยกับแผ่นกระดาษเสีย 0, 1, 2, 3, 4, 5 ในการห่อ จงเติมในช่องว่างในตารางขาดทุน

a	θ	0	1	2	3	4	5
a_1		.25					
a_2		−.75		.25			
a_3		−.25	−.15		.05		

(ข) สมมติว่าแบบจำลองสำหรับกระบวนการที่แต่ละแผ่นกระดาษมีการแจกแจงแบบเบอร์นูลลี พร้อมด้วยพารามิเตอร์ p = 0.05 สภาวะธรรมชาติของ prior distribution เป็นแบบทวินามพร้อมด้วยพารามิเตอร์ n = 5 และ p = 0.05 ถ้า θ เป็นจำนวนแผ่นกระดาษเสียแล้ว

$$P(\theta = k) = P\theta(k) = \binom{5}{k} (0.5)^{k} (.95)^{5-k}$$

จงคำนวณหากระบวนการของเบย์ส (ก่อนการทดสอบแผ่นกระดาษ)

(a) ภายหลังลงมือทดสอบแผ่นกระดาษจงคำนวณหากระบวนการของเบย์ส ถ้าหากว่าการพิมพ์ดี ถ้าหากว่าการพิมพ์เลว