บทที่ 3. ตัวแปรเชิงสุ่มและการแจกแจงน่าจะเป็น

THE MOST IMPORTANT QUESTIONS OF LIFE ARE, FOR THE MOST PART, REALLY ONLY PROBLEMS OF PROBABILITY.

— Pierre Simon de Laplace

การทดลองเชิงสุ่มบางประเภทให้ผลทดลองเป็นตัวเลข (Numbers) การทดลองประเภทนี้มักจะมีการวัดหรือนับประมาณบางอย่าง เช่น ถ้าจะนับจำนวนลูกค้าที่เข้ามาในธนาคารระหว่าง 9.00-10.00 น. ผลลัพธ์จากการนับ (ซึ่งเป็นตัวเลข) เป็นผลจากการทำการทดลองเชิงสุ่ม เพราะเราไม่ทราบล่วงหน้าก่อนการนับว่าจะมีลูกค้าเข้ามากี่คน ในกรณีนี้ กลุ่มผลทดลองจะเป็น $S = \{0, 1, 2, ...\}$

ตัวเลขอาจจะได้มาจากการวัดเช่น การวัดเวลาที่หลอดไฟฟ้าตั้งแต่เริ่มเปิดจนถึงหลอดเสีย (หรืออายุของหลอด) ว่าเป็นเท่าใด การนับหรือวัดซึ่งเป็นการทดลองเชิงสุ่มทำนองนี้เราเรียกว่า การนับเชิงสุ่ม (Random Counting) หรือการวัดเชิงสุ่ม (Random Measurement) ก่อนที่จะทำการนับหรือวัดแล้วเสร็จ ตัวเลขที่จะได้เป็นผลลัพธ์เรามักจะแทนด้วย X และเรียก X ว่า "ตัวแปรเชิงสุ่ม" (Random, Chance, or Stochastic Variable) หรือตัวอื่น ๆ เช่น Y, Z

จะสังเกตว่า X เป็นแต่เพียงสัญญลักษณ์แทนตัวเลขที่ยังไม่เป็นค่าใดค่าหนึ่งจนกว่าจะทำการนับหรือวัดเชิงสุ่มแล้วเสร็จ X จึงเป็นเพียงสัญญลักษณ์หนึ่งซึ่งแปรไปและขึ้นกับว่าผลลัพธ์ของการนับเชิงสุ่มจะเป็นอย่างไร ในเมื่อทำการนับเชิงสุ่มเสร็จแล้ว ตัวแปรเชิงสุ่ม X จึงจะมีผลลัพธ์เป็นค่าใดค่าหนึ่งโดยเฉพาะ

การศึกษาถึงตัวแปรสุ่มนี้ เรามิได้มุ่งสนใจดูว่า X จะออกผลลัพธ์เป็นอะไร ความสนใจจะมุ่งอยู่ที่สถานะการณ์ก่อนจะทำการนับแล้วเสร็จ และการคำนวณหาความน่าจะเป็นที่ X ออกเป็นค่าใดค่าหนึ่ง เช่น X หรือความน่าจะเป็นที่ X จะออกผลลัพธ์อยู่ในช่วงหนึ่ง เช่น ช่วง (a, b) เราแทนความน่าจะเป็นเหล่านี้ด้วย $P(X = x)$ และ $P(a < X < b)$

ในแง่คณิตศาสตร์เรามองตัวแปรเชิงสุ่มในเชิงฟังก์ชันทางคณิตศาสตร์ประเภทหนึ่งได้ดังนิยามต่อไปนี้

นิยาม ตัวแปรเชิงสุ่ม X เป็นฟังก์ชันค่าจริง (Real-valued Function) ในตัวแบบน่าจะเป็น (S, P) โดยที่มันจะกำหนดแต่ละผลทดลองด้วยเลขจำนวนจริง ซึ่งเรียกว่าค่าของตัวแปรเชิงสุ่มของผลทดลองนั้น

นั่นคือตัวแปรเชิงสุ่มเป็นฟังก์ชันค่าจริงที่มีโดเมนเป็นกลุ่มผลทดลองของการทดลองเชิงสุ่มและมีพิสัยเป็นเซทของเลขจำนวนจริง

ดังนั้น ตัวแปรเชิงสุ่มจะให้ค่าหนึ่งต่อแต่ละจุดหรือผลทดลอง เช่น ถ้ากำหนดจุด

ของกลุ่มผลทดลอง S แล้ว $X(s)$ จะเป็นเลขจำนวนจริงจำนวนหนึ่ง ถ้าให้ X เป็นตัวแปรเชิงสุ่มที่มีค่า $x_1, x_2, x_3, \ldots$ นั่นคือ ค่าหรือพิสัยของ X เป็น $x_1, x_2, x_3, \ldots$ สำหรับแต่ละ i เราเขียน $X = x$ แทนเหตุการณ์ที่ประกอบด้วยจุดที่แปลงไปเป็น x_i ด้วย x นั่นคือ

$$X = x = \{ s \in S \mid X(s) = x_i \}$$

เราจะเขียนความน่าจะเป็นของเหตุการณ์ $X = x_i$ ด้วย $f(x_i)$ ซึ่งมีความหมายเท่ากับ $P(X = x_i)$

ตัวอย่าง ในการทดลองโดยการหยิบหลอดไฟมา 2 หลอด จากกล่องหนึ่ง ให้ X เป็นตัวแปรเชิงสุ่มที่แทนจำนวนหลอดไฟที่ดี ดังนั้น กลุ่มผลทดลอง S = {GG, GD, DG, DD} จะเป็นโดเมนของ X และเซทของค่าที่แทนจำนวนหลอดดี {0 1 2} จะเป็นพิสัยของ X หรือเซทของค่าที่ X จะมีได้ (D และ G เป็นหลอดเสียและดีตามลำดับ)

ตัวแปรเชิงสุ่มจะแปลงแต่ละจุดในโดเมนหรือกลุ่มผลทดลองไปยังจุดหนึ่งในพิสัย

GG	GD DG	DD	โดเมนของ X
↓	↘ ↙	↓	
2	1	0	พิสัยของ X

หรือบางทีเราจะเขียนเป็นเชิงคณิตศาสตร์ได้ ดังนี้

$$X(GG) = 2 \qquad X(GD) = 1$$
$$X(DG) = 1 \qquad X(DD) = 0$$

ซึ่งหมายความว่า ถ้าผลทดลองเป็น GG ให้ X มีค่าเป็น 2 ถ้าหากเป็น GD ให้เป็น 1, และต่อ ๆ ไป

3.1 ประเภทของตัวแปรเชิงสุ่ม (Classification of Random Variable)

เนื่องจากตัวแปรเชิงสุ่มเป็นฟังก์ชั่นที่นิยามหรือกำหนดในกลุ่มผลทดลองและโดยที่กลุ่มผลทดลองมีอยู่ 2 ประเภท เราจึงแยกตัวแปรเชิงสุ่มตามกลุ่มผลทดลองได้เป็น 2 ประเภทเหมือนกัน คือ

(1) ตัวแปรเชิงสุ่มไม่ต่อเนื่อง (Discrete Random Variable) เป็นตัวแปรเชิงสุ่มที่นิยามในกลุ่มผลทดลองไม่ต่อเนื่อง (Discrete Sample Space) ตัวแปรเชิงสุ่มประเภทนี้ มักจะเกิดจากการนับเชิงสุ่ม (Random Counting) ตัวอย่างเช่น จำนวนอุบัติเหตุที่เกิดขึ้นแต่ละปีบนถนนสาย 71 จะเป็นตัวแปรเชิงสุ่มไม่ต่อเนื่องที่มีค่าหนึ่งค่าใดในเซทของ 0 1, 2, 3, ... นั่นคือ ผลลัพธ์ของตัวแปรเชิงสุ่มจะเป็นเลขตัวเต็ม (Real integer) แต่ในกรณีอื่น ถึงแม้ว่าเป็นตัวแปรเชิงสุ่มไม่ต่อเนื่อง ผลลัพธ์อาจจะไม่เป็นเลขตัวเต็มก็ได้ เช่น กรณีที่กลุ่มผลทดลองของ X มีตัวเลขที่เป็นไปได้เพียง 5 จำนวน คือ

$$S = \{3.50, 3.78, 15.12, 17.25, 25.30\}$$

X ก็ยังคงเป็นตัวแปรเชิงสุ่มไม่ต่อเนื่องเพราะ กลุ่มทดลองเป็นแบบไม่ต่อเนื่อง **กรณีนี้ผลลัพธ์**ของ X ไม่เป็นตัวเลขตัวเต็ม

(2) ตัวแปรเชิงสุ่มต่อเนื่อง (Continuous Random Variable) เป็นตัวแปรเชิงสุ่มที่นิยามในกลุ่มผลทดลองต่อเนื่อง (Continuous Sample Space) ตัวอย่างเช่นน้ำหนักของน้ำอัดลมที่บรรจุโดยเครื่องจักรจะเป็นตัวแปรเชิงสุ่มต่อเนื่องซึ่งจะมีค่าหนึ่งค่าใดในช่วง (a, b) ในเมื่อ a, b เป็นค่าจริงที่ $a<b$

ถ้าจะพิจารณาในค่าของตัวแปรเชิงสุ่มหรือกฎในการกำหนดเลขจำนวนจริงให้แก่ผลทดลองในกลุ่มผลทดลอง แล้วเราสามารถแยกออกตัวแปรเชิงสุ่ม ออกเป็น 4 ประเภทตามระดับของการวัด (Level of Measmement) โดยที่การวัดเป็นการกำหนดตัวเลขให้แก่สิ่งของหรือเหตุการณ์ตามกฎต่าง ๆ และกฎต่าง ๆ กันนี้เองจะให้ระดับของการวัดต่าง ๆ กันด้วย

(1) ตัวแปรเชิงสุ่มแบบนามบัญญัติ (Nominal or Categorical Random Variables) ค่าของตัวแปรเชิงสุ่มประเภทนี้ใช้ขนานชื่อให้กับผลทดลองหรือเหตุการณ์ต่าง ๆ เช่น หญิง ชาย เป็นต้น ตัวเลขที่เป็นค่าของตัวแปรเชิงสุ่มประเภทนี้จะมีคุณค่า (value) เหมือนกัน และอาจใช้สลับสับเปลี่ยนหรือทดแทนกันได้ เช่น 0 ให้เป็นชาย และ 1 เป็นหญิง หรือจะให้ 1 เป็นชาย และ 0 เป็นหญิงก็ได้

(2) ตัวแปรเชิงสุ่มแบบเรียงอันดับ (Ordinal of Rank Random Variables) ค่าของตัวแปรเชิงสุ่มประเภทนี้ใช้เรียง หรือจัดอันดับผลทดลองในสกุลเดียวกันให้ลดหลั่นกันเป็นขั้น ๆ ตามปริมาณและคุณภาพมากน้อย เช่นระดับรายได้ต่ำกว่า, และสูง ก็กำหนดตัวเลขเป็น 1, 2, 3 เป็นต้น.

(3) ตัวแปรเชิงสุ่มแบบอันตรภาค (Interval Random Variables) ค่าของตัวแปรเชิงสุ่มประเภทนี้ใช้แบ่งตัวกำหนด ผลทดลองในสกุลเดียวกันออกเป็นประเภท ๆ หรือเป็นจังหวะ เป็นช่วง ๆ ขั้น ๆ ที่มีขนาดใหญ่เท่า ๆ กัน อย่างไรก็ตามตัวเลขที่กำหนดนั้นถ้าเป็น 0 ก็ไม่ได้หมายความเป็นศูนย์แท้จริงศูนย์อนันต์ (Absolute zero) แต่เป็นศูนย์นิยมหรือศูนย์สัมพัทธ์ (Relative or arbitrary zero) เท่านั้น และจำนวนเท่าของค่าของตัวแปรเชิงสุ่มจะไม่เป็นจริง เช่นค่าของตัวแปรที่เป็นคะแนนสอบ ถ้าได้ 0 ก็ไม่ได้หมายความว่าไม่มีความรู้ หรือได้ 50 ก็ไม่ใช่มีความรู้เป็น 2 เท่าของ 25 เป็นต้น

(4) ถ้าแปรเชิงสุ่มแบบอัตราส่วน (Ratio Random Variables) ค่าของตัวแปรประเภทนี้เป็นจำนวนเลขที่สมบูรณ์ที่สุด นั่นคือมีศูนย์แท้จริงศูนย์อนันต์ มีหน่วยเป็นขนาดโตเท่า ๆ กัน และเรียงขึ้นลงตามลำดับสม่ำเสมอกัน และจำนวนเท่าของค่าจะเป็นจริง เช่นค่าของแปรเชิงสุ่มที่ใช้แทนน้ำหนัก, ความยาว, ความหนาแน่น เป็นต้น

เราพอจะสรุปลักษณะที่สำคัญของตัวแปรเชิงสุ่มทั้ง 4 ประเภท ได้ดังตารางต่อไปนี้

ลักษณะที่สำคัญ

ตัวแปรเชิงสุ่ม	ก. อันดับ	ข. ระยะยาว	ค. จุดกำเนิด
นามบัญญัติ	x	x	x
เรียงอันดับ	✓	x	x
อันตรภาค	✓	✓	x
อัตราส่วน	✓	✓	✓

ก. อันดับ (Sequence or Order) ของจำนวนมีความหมายหรือไม่?

ข. ระยะทาง (Distance) ระหว่างสองจำนวนมี[illegible]หรือไม่?

ค. จุดกำเนิด (Origin) มีความหมายเดียว (Unique) หรือไม่?

ตัวอย่าง ในสัปดาห์หนึ่งเราอาจจะใช้เครื่องมือสำหรับเขียนหลายชนิด เช่น ดินสอ ปากกาหมึกซึม ปากกาลูกลื่น ปากกาแมจิก เป็นต้น ตัวแปรเชิงสุ่มแบบใดที่จะใช้เหตุการณ์ต่อไปนี้

1. ให้เบอร์เครื่องเขียน
2. เรียงลำดับเครื่องเขียนตามความชอบ
3. วัดอุณหภูมิต่ำสุดที่เครื่องเขียนจะใช้งานได้
4. วัดน้ำหนักเครื่องเขียน

ตัวแปรเชิงสุ่มที่ใช้ในแต่ละเหตุการณ์จะต่างกัน นั่นคือเป็นตัวแปรเชิงสุ่มแบบนามบัญญัติ เรียงอันดับ อันตรภาค และอัตราส่วน

3.2 ตัวแปรเชิงสุ่มไม่ต่อเนื่องและฟังก์ชั่นมวลน่าจะเป็น (Discrete Random Variable and Probability Mass Function P.M.F.)

ลักษณะที่สำคัญของตัวแปรเชิงสุ่มไม่ต่อเนื่องก็คือ

(1) ค่าของตัวแปรที่เป็นไปได้

(2) ความน่าจะเป็นที่เกี่ยวพัน (associated) กับแต่ละค่าที่เป็นไปได้

จากการที่เราทราบว่า ความน่าจะเป็นของแต่ละผลทดลองในกลุ่มผลทดลองเป็นเท่าใด เราก็สามารถกำหนดความน่าจะเป็นที่ตัวแปรเชิงสุ่มจะออกผล (หรือมีค่า) เป็นค่าใด ๆ ก็ได้ **กฎ**หรือวิธีการที่กำหนดความน่าจะเป็นให้แก่ค่าของตัวแปรเชิงสุ่มไม่ต่อเนื่องเราเรียกว่า**ฟังก์ชั่นมวลน่าจะเป็น** (Probability Mass Function) หรือการแจกแจงน่าจะเป็น **(Probability Distribution)** ดังนิยามต่อไปนี้

นิยาม กำหนดให้ x เป็นตัวแปรเชิงสุ่มไม่ต่อเนื่องที่มีค่าเป็น $x_1, x_2, x_3, \ldots$ ฟังก์ชั่น f

จะเรียกว่าฟังก์ชั่นมวลน่าจะเป็น (Probability Mass Function, PMF) ของ X ถ้า f กำหนดไว้ดังนี้ $f(x_i) = P(X=x_i)$, $i=1,2,3,\ldots\ldots$ และ $f(x_i)$ จะต้องมีคุณสมบัติดังนี้

(1) $0 \leq f(x_i) \leq 1$

(2) $\sum_i f(x_i) = 1$

ตัวอย่าง การทดลองอย่างหนึ่งโดยการเลือกสุ่มชิ้นส่วน 3 ชิ้นจากขบวนการ ผลิตซึ่งมีจำนวนของดีและของเสียเท่า ๆ กัน

ให้ตัวแปรเชิงสุ่ม x เป็นจำนวนชิ้นส่วนที่ไม่ดี ที่สุ่มมาแล้วจุดหรือสมาชิกของกลุ่มผลทดลองสำหรับการทดลองนี้กับความน่าจะเป็นของแต่ละสมาชิกจะเป็นตารางต่อไปนี้

ผลทดลอง (s_i)	GGG	GGD	GDG	DGG	GDD	DGD	DDG	DDD
ความน่าจะเป็น $P(s_i)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
ค่าของ X หรือ x_i	0	1	1	1	2	2	2	3

เหตุการณ์สำหรับ $x = 0$ มี 1 จุด (s_1) และมีความน่าจะเป็น $= \frac{1}{8}$

เหตุการณ์สำหรับ $X = 1$ มี 3 จุด (s_2, s_3, s_4) ดังนั้นความน่าจะเป็นจึงเท่ากับ

$$\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{3}{8}$$

เหตุการณ์สำหรับ $X = 2$ มี 3 จุด (s_5, s_6, s_7) ความน่าจะเป็นจึงเท่ากับ

$$\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{3}{8}$$

เหตุการณ์สำหรับ $X = 3$ มี 1 จุด (s_8) ความน่าจะเป็นจึงเท่ากับ 1/8

ค่าที่เป็นไปได้ของ X และความน่าจะเป็นของมันดังตารางข้างล่างหรือเซทของคู่ลำดับ $(x, f(x))$ นี้เป็นการแจกแจงน่าจะเป็นของ X

ค่าของตัวแปรเชิงสุ่ม $X(x_i)$	0	1	2	3
ความน่าจะเป็น $f(x_i)$	1/8	3/8	3/8	1/8

สำหรับการแจกแจงความน่าจะเป็นนั้นเราสามารถเขียนให้อยู่ในรูปสมการหรือกราฟก็ได้ ซึ่งมันจะทำหน้าที่สรุปตารางข้างบนนี้นั่นเอง

1. สมการหรือสูตร (Equation or Formula)

$$f(x) = \begin{cases} 1/8, & x = 0, 3 \\ 3/8, & x = 1, 2 \end{cases}$$

หรือ $f(x) = \binom{3}{x}(1/2)^3, x = 0, 1, 2, 3$

2. กราฟ (Graph)

ตัวอย่าง นักเศรษฐศาสตร์ประจำบริษัทกลุ่มหนึ่งได้พัฒนาการแจกแจงความน่าจะเป็นเชิงจิตวิสัย (Subjective Probability Distribution) ของการเปลี่ยนแปลงที่เกิดขึ้นในปีต่อไปของผลิตภัณฑ์ประชาชาติรวม (GNP) เขาได้สร้างประเภท (Categories) ของการเปลี่ยนแปลงนั้นเป็น 5 ประเภทและนิยามตัวแปรเชิงสุ่มโดยการให้ตัวเลขบ่งถึงแต่ละประเภทดังนี้

การเปลี่ยนแปลงของ GNP	จำนวนเลขที่กำหนด
ลดมากกว่า 5%	- 2
ลด 5% หรือน้อยกว่า	- 1
ไม่เปลี่ยนแปลง	0
ขึ้น 5% หรือน้อยกว่า	+ 1
ขึ้นกว่า 5%	+ 2

จากข้อมูลข่าวสารที่เอื้ออำนวยให้เขาจึงกำหนดความน่าจะเป็นให้แก่แต่ละเหตุการณ์ที่เป็นไปได้เหล่านี้ ดังตาราง

x	-2	-1	0	1	2
f(x)	0.1	0.1	0.2	0.4	0.2

3.3 ตัวแปรเชิงสุ่มต่อเนื่องและฟังก์ชั่นหนาแน่นน่าจะเป็น (Probability Density Function, PDF)

ในกรณีที่ชุดของตัวเลขที่เป็นพิสัยของตัวแปรเชิงสุ่ม มีจำนวนตัวเลขมากเป็นอนันต์จนนับไม่ได้ (Uncountably Infinite) หรือกรณีที่กลุ่มผลทดของ X เป็นแบบต่อเนื่องนั่นเอง ตัวแปรเชิงสุ่มนี้จะเป็นผลลัพธ์ของ X ในกลุ่มผลทดลองของ X จะมีจำนวนมากเป็นอนันต์และนับไม่ได้ด้วย ตัวแปรเชิงสุ่มประเภทนี้ส่วนมากเกิดจากการวัดเชิงสุ่ม เช่น การทดลองเกี่ยวกับการวัดเวลาเดินทางจากบ้านมายังมหาวิทยาลัย เวลาที่วัดได้เป็นตัวแปรเชิงสุ่ม X ซึ่งจะมีค่าใดเป็นจุดทศนิยมกี่ตัวก็ได้ในช่วง $(0, \infty)$ ในช่วงนี้มีจำนวนเลขมากมายเป็นอนันต์นับไม่ได้

เมื่อกลุ่มทดลองเป็นแบบไม่ต่อเนื่องเราจะกำหนดความน่าจะเป็นที่ไม่เป็นศูนย์ให้แก่แต่ละค่าของตัวแปรเชิงสุ่มได้ ถึงแม้ค่าของตัวแปรเชิงสุ่มจะเป็นอนันต์ที่นับได้ (Countable infinite) และความน่าจะเป็นที่กำหนดให้นั้นรวมกันเท่ากับหนึ่ง ส่วนกลุ่มผลทดลองที่ต่อเนื่องนี้เป็นไปไม่ได้ ถ้าเราพยายามสร้างเหตุการณ์ง่าย ๆ (Simple Events) ที่ประกอบด้วยผลทดลองเพียงจุดเดียว และกำหนดความน่าจะเป็นให้แต่ละเหตุการณ์ง่าย ๆ เหล่านี้เพื่อให้ผลรวมของความน่าจะเป็นเท่ากับ 1 แล้วเราจะต้องพบกับงานที่ไม่สามารถทำได้ ดังนั้นจึงหันมาสนใจเหตุการณ์ที่ประกอบด้วยผลทดลองมาก ๆ จุด (Compound Event) คือเหตุการณ์ที่ตัวแปรเชิงสุ่มมีค่าในช่วงใดช่วง

หนึ่ง มากกว่าที่จะสนใจเป็นจุด ๆ ไป เราสามารถนิยามฟังก์ชั่นหนาแน่นน่าจะเป็นได้เป็นดังนี้

นิยาม ให้ X เป็นตัวแปรเชิงสุ่มต่อเนื่องที่มีพิสัยเป็นเซทของเลขจำนวนจริงที่อยู่ในช่วง (a,b) a < b เรานิยามฟังก์ชั่น f ที่เรียกว่าฟังก์ชั่นหนาแน่นน่าจะเป็น (PDF, Probability Density Function) ถ้า f(x) มีคุณสมบัติดังต่อไปนี้

1. $f(x) \geq 0 \; ; \quad a \leq x \leq b$
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$
3. ความน่าจะเป็นที่ X มีค่าอยู่ในช่วง $x,\ x+dx = f(x)\, dx$

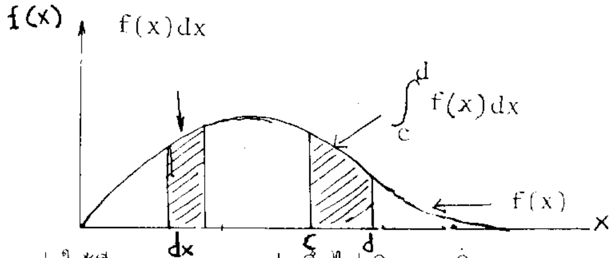
   นั่นคือ $P[x \leq X \leq x+dx] = f(x)\, dx$

หากสามารถหา f(x) ของตัวแปรเชิงสุ่มของ X ได้ เราก็จะสามารถกำหนดความน่าจะเป็นที่ X จะมีค่าในช่วง [x, x+dx] เป็น f(x) dx ซึ่งจะเท่ากับพื้นที่ภายใต้ฟังก์ชั่นหรือโค้ง f(x) ระหว่างจุด x กับ x+dx ดังนั้นเราสามารถคำนวณความน่าจะเป็นที่ X จะมีค่าอยู่ในช่วง c,d ได้ด้วย

$$P[c \leq x \leq d] = \int_c^d f(x)\, dx$$

ซึ่งเท่ากับพื้นที่ภายใต้ฟังก์ชั่น f(x) ระหว่างจุด c กับ d

PDF สามารถแสดงได้เป็นกราฟที่ต่อเนื่องดังนี้



ตัวอย่าง ตัดเหล็กแท่งให้มีความยาวแบบสุ่มซึ่งไม่เกิน 12 นิ้ว

P.D.F. สำหรับตัวแปรเชิงสุ่ม x ที่แทนความยาวของเหล็กแท่งซึ่งเป็นนิ้วจะหาได้ดังนี้

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^{12} a\,dx = 1$$

$$ax \Big|_0^{12} = 1 \; ; \; a = 1/12$$

ดังนั้น $f(x) = 1/12 \; ; \; 0 < x < 12$

ถ้าจะเสนอ PDF ของ x หรือ $f(x) = 1/12, 0 < x < 12$ ในรูปตารางเราต้องให้ค่าของตัวแปรเชิงสุ่มเป็นช่วง ๆ ดังนี้

x	f(x)
$0 < x < 4$	1/3
$4 < x < 8$	1/3
$8 < x < 12$	1/3

3.4 ฟังก์ชั่นแจกแจงสะสม (CDF, Cumulative Distribution Function)

ตัวแปรเชิงสุ่มที่ผ่านมาเราเสนอหรือบรรยายคุณลักษณะของมันด้วย PMF.และ PDF เราอาจจะใช้ฟังก์ชั่นแจกแจงสะสม (C.D.F.) มาบรรยายคุณลักษณะตัวแปรเชิงสุ่มได้ ซึ่งเราจะนิยาม CDF ดังนี้

นิยาม ให้ X เป็นตัวแปรเชิงสุ่ม ฟังก์ชั่น F จะเรียกว่าฟังก์ชั่นแจกแจงสะสม (Comulative Distribution Function, CDF) ถ้า F(x) กำหนดไว้ดังนี้

$$F(x) = P[X \le x] = \begin{cases} \sum_{X \le x} f(x); & x \text{ ไม่ต่อเนื่อง} \\ \int_{-\infty}^{x} f(x)dx; & x \text{ ต่อเนื่อง} \end{cases}$$

และ F(x) ต้องมีคุณสมบัติดังต่อไปนี้

(1) F(x) จะเป็นฟังก์ชั่นที่มีค่าเพิ่มขึ้นเมื่อ x มีค่าใหญ่ขึ้น ซึ่งเรียกว่าเป็นฟังก์ชั่งขยาย (Nondecreasing function) นั่นคือ ถ้า $x_1 < x_2$ แล้ว $F(x_1) < F(x_2)$

(2) $F(+\infty) = 1$ และ $F(-\infty) = 0$ หรือ $\lim_{x \to -\infty} F(x) = 0$ และ $\lim_{x \to \infty} F(x) = 1$ นั่นคือ $0 \le F(x) \le 1$

(3) ถ้า X เป็นตัวแปรเชิงสุ่มต่อเนื่อง $f(x) = \frac{d\,F(x)}{dx}$ และถ้า x เป็นตัวแปรเชิงสุ่มไม่ต่อเนื่องที่มีค่าเป็น $x_1, x_2, x_3, \ldots$ โดยที่ $x_1 < x_2 < x_3 \ldots$ แล้ว $f(x_i) = F(x_i) - F(x_{i-1})$

(4) ถ้า x เป็นตัวแปรเชิงสุ่มไม่ต่อเนื่อง F(x) จะเป็นฟังก์ชั่นขั้นบันได (Step function) ซึ่งเป็นฟังก์ชั่นที่ต่อเนื่องทางขวา (Continuous on the right) แต่ไม่ต่อเนื่องทางซ้ายและขั้นบันไดนั้นเกิดขึ้นที่จุด $X = x_i$ พอดี

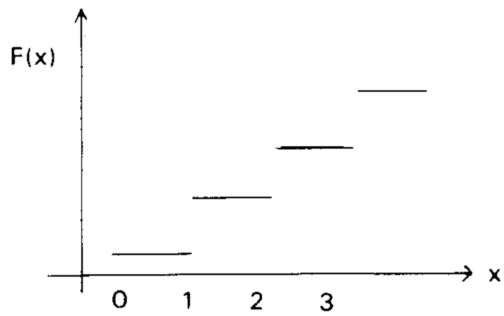
ถ้า X เป็นตัวแปรเชิงสุ่มต่อเนื่อง F(x) จะเป็นฟังก์ชั่นที่ต่อเนื่อง

ตัวอย่าง ให้ X เป็นตัวแปรเชิงสุ่มที่ฟังก์ชั่นมวลน่าจะเป็น (PMF) ดังนี้

x	0	1	2	3
f(x)	1/8	3/8	3/8	1/8

เราสามารถหาฟังก์ชั่นแจกแจงสะสม (CDF) ของ x ได้ในรูปของตารางหรือกราฟดังนี้

ค่าของตัวแปรเชิงสุ่ม x	ความน่าจะเป็น F(x)
$x < 0$	0
$x \le 0$	1/8
$x \le 1$	4/8
$x \le 2$	7/8
$x \le 3$	1



ตัวอย่าง ฟังก์ชันหนาแน่นน่าจะเป็น (PDF) ของ X กำหนดไว้ดังนี้

$$f(x) = 1/(b-a);\ a < x < b$$

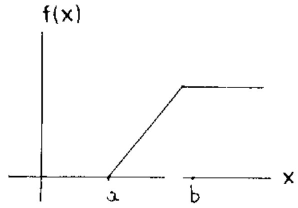
$$\text{จาก}F(\infty) = \int_{-\infty}^{\infty} f(x)dx$$

$$= \int_{-\infty}^{\infty} 0dx + \int_{a}^{b} 1/(b-a)dx + \int_{b}^{\infty} 0dx$$

$$= 1$$

เราจะได้ F(x) และกราฟของ f(x) ดังนี้

$$F(x) = \begin{cases} 0 \ ; & x < a \\ (x-a)/(b-a); & a < x < b \\ 1 \ ; & x > b \end{cases}$$



3.5 **มาตราวัดสรุปของตัวแปรเชิงสุ่ม** (Summary Measures of Random Variables)

เราได้พิจารณาคุณลักษณะของตัวแปรเชิงสุ่มในแง่ฟังก์ชั่นน่าจะเป็นและฟังก์ชั่นแจกแจงสะสมมาแล้ว **ต่อไปเราจะพิจารณาคุณลักษณะของตัวแปรเชิงสุ่มด้วยดัชนีหรือปริมาณ**ที่เป็นจำนวนเลขเพียงจำนวนเดียว ปริมาณนี้เราเรียกว่า มาตรประชากร หรือ พารามิเตอร์ (Parater) ของการแจกแจงของตัวแปรเชิงสุ่มสำหรับมาตราวัดสรุปตัวแปรเชิงสุ่มที่เรารู้จักกันดีก็มีดังต่อไปนี้

3.5.1 **มาตราวัดค่ากลางและมาตราวัดตำแหน่ง** (Measures of Location and Position)

ก. มาตราวัดค่ากลาง เป็นดัชนีหรือปริมาณจำนวนหนึ่ง (พารามิเตอร์) มีคุณลักษณะของตัวแปรเชิงสุ่มโดยค่ากลาง ๆ (Expected, typical or normative value) เรามีวิธีวัดค่าต่าง ๆ ดังนี้

(1) **กึ่งพิสัย** (Midrange) เป็นมาตราวัดที่หยาบที่สุด หาได้รวดเร็วดังนิยามต่อไปนี้

นิยาม ถ้า x เป็นตัวแปรเชิงสุ่มโดยมี x_i เป็นค่าของตัวแปรเชิงสุ่มแล้วกึ่งพิสัยจะเป็น

$$m_r = \frac{1}{2}(Max(x_i) + Min(x_i))$$

ในเมื่อ $Max(x_i)$ และ $Min(x_i)$ เป็นค่าของตัวแปรที่มากที่สุดและต่ำสุดตามลำดับ

ตัวอย่าง ถ้า X มีการแจกแจงดังนี้

x	1	2	3	4
f(x)	0.1	0.4	0.3	0.2

แล้ว $m_r = \frac{1}{2}(4+1) = 2.5$

(2) **ฐานนิยม** (Mode) เป็นมาตราวัดค่าเฉลี่ยหรือค่ากลางซึ่งกำหนดไว้ว่า

นิยาม ถ้า X เป็นตัวแปรเชิงสุ่ม x_o จะเป็นฐานนิยมของ X ถ้า x_o เป็นค่าที่ทำให้ฟังก์ชั่นของ X หรือ f(x) มีค่าสูงสุด

นั่นคือ ถ้า X เป็นตัวแปรเชิงสุ่มไม่ต่อเนื่อง ค่าของ X ที่มีความน่าจะเป็น

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

และถ้า X เป็นตัวแปรเชิงสุ่มต่อเนื่อง x_o จะเป็นฐานนิยมของ X ถ้าสอดคล้องกับสมการ

(a) $f'(x)\big|_{x=x_0} = 0$

(b) $f''(x)\big|_{x=x_o} < 0$

ฐานนิยมของตัวแปรเชิงสุ่มตัวหนึ่งจะหาได้เสมอ (exist) แต่ไม่จำเป็นต้องมีตัวเดียว (Unique) แต่ถ้ามีหลายตัวก็ไม่มีประโยชน์ในการเป็นมาตราวัดสรุป

เราสนใจฐานนิยมเมื่อต้องการทราบว่าเหตุการณ์ไหนหรือค่าไหนของตัวแปรงเชิงสุ่มมีโอกาสเกิดขึ้นมากที่สุด เช่นสนใจว่าจำนวนเด็กในครอบครัวเป็นอย่างไร เราใช้ฐานนิยมดีกว่าอย่างอื่น ตัวแปรเชิงสุ่มแบบนามบัญญัติ หรือเรียงอันดับจะสรุปด้วยฐานนิยม สำหรับตัวแปรเชิงสุ่มอย่างอื่นก็สรุปด้วยฐานนิยมได้ แต่มักมีมาตราวัดอื่น ๆ ที่ดีกว่า.

ตัวอย่าง ก. ถ้าตัวแปรเชิงสุ่ม x ที่แทนจำนวนบุตรในครอบครัวของผู้จบปริญญา

X	1	2	3
f(x)	.5	.2	.3

โดยที่ f(1) สูงสุดดังนั้น $x_o = 1$ จึงเป็นฐานนิยมซึ่งใช้วัดค่ากลางได้ดีพอสมควร

ข. ถ้า x แจกแจงดังนี้

$f(x) = e^{-x}, \; x < o$

เนื่องจาก f(o) มากสุด ดังนั้น $x_o = o$ จึงเป็นฐานนิยม แต่ใช้วัดค่ากลางได้ไม่ดี.

ค. ถ้า x แจกจงดังนี้

$$f(x) = 1/\sqrt{2\pi}\, e^{-x^2/2} \quad ; \quad -\infty < x < \infty$$

$$f'(x) = 1/\sqrt{2\pi}\, e^{-x^2/2}\,(-x)$$

เมื่อ $f'(x) = o$ เราได้ $x = -\infty, 0, +\infty$ แต่ $f''(o) < 0$

ดังนั้น $x_o = 0$ เป็นฐานนิยมและใช้ได้ดี

(3) **มัธยฐาน** (Median) เป็นค่าที่แบ่งการแจกแจงของตัวแปรเชิงสุ่มออกเป็นสองส่วนเท่า ๆ กัน ดังนิยามต่อไปนี้

นิยาม ถ้า x_m เป็นค่าของตัวแปรเชิงสุ่ม x ที่ทำให้

$P(X \le x_m) \geq 1/2$ และ $P(X \geq x_m) \geq 1/2$ แล้ว x_m จะเรียกว่ามัธยฐานของ x

กรณี X เป็นตัวแปรเชิงสุ่มต่อเนื่องเราจะได้ว่า $F(x_m) = P(X \le x_m) = \frac{1}{2}$ นั่นคือ เราหา x_m จากสมการ $\int_{-\infty}^{x_m} f(x)dx = 1/2$

ตัวอย่าง ก. กำหนดฟังก์ชันน่าจะเป็นดังนี้

x	0	1	2	3
f(x)	0.1	0.2	0.3	0.4

$x_m = 2$ เป็นมัธยฐานเพราะ $P(x \leq 2) = 0.6 > 0.5$ และ $P(x \geq 2) = 0.7 > 0.5$

เราสามารถหาได้โดยวิธีเทียบ (Process of Interpolation) ดังนี้

x	0	1	2	3
F(x)	0.1	0.3	0.6	1

$$x_m = 1.5 + \frac{0.5-0.3}{0.6-0.3} = 2.17 \text{ หรือ}$$

$$x_m = 1.5 + \frac{0.2}{0.3}(1) = 2.17$$

จากนิยามและตัวอย่างเราจะเห็นได้ว่ามัธยฐานของตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องจะมีเสมอ แต่ไม่จำเป็นว่าต้องมีตัวเดียว

ข. กำหนดให้ฟังก์ชั่นหนาแน่นน่าจะเป็นไว้ดังนี้

$f(x) = 2x/3;\ 1 < x < 2$

หามัธยฐาน x_m จากสมการ

$$\int_0^{x_m} (2x/3)dx = 0.5\ ,\ (x^2)/3\Big|_1^{x_m} = 0.5\ ,\quad x_m = 1.58$$

มัธยฐานมักจะใช้สรุปตัวแปรเชิงสุ่มแบบเรียงอันดับ หรืออาจจะใช้กับตัวแปรเชิงสุ่มแบบอันตรภาคและอัตราส่วน

ในการศึกษาการแจกแจงของรายได้ของครอบครัว โดยเฉพาะในประเทศด้อยพัฒนานั้น มัธยฐานจะอธิบายหรือแสดงคุณลักษณะของการแจกแจงรายได้ดีกว่าอย่างอื่น เพราะว่าครอบครัวที่มีรายได้มากมีอยู่น้อย

(4) ค่าคาดหวัง (Mean of Expected Value) เป็นส่วนเฉลี่ยเลขคณิตแบบถ่วงน้ำหนัก (Weighted average mean) และเป็นมาตราวัดค่ากลางที่มนุษย์รู้จักและใช้กันมากมานานแล้ว Lambert Quetelet (1796-1874) นักดาราศาสตร์และสถิติชาวเบลเยี่ยมเป็นผู้พัฒนาขึ้นมา

นิยาม ถ้า X เป็นตัวแปรเชิงสุ่ม ค่าคาดหวังของ X ซึ่งแทนด้วย E(x) กำหนดไว้ดังนี้

$$E(X) = \sum x\,f(x)\ ;\ X \text{ เป็นตัวแปรเชิงสุ่มไม่ต่อเนื่อง}$$

$$= \int_{\forall x} x f(x)dx\ ;\ X \text{ เป็นตัวแปรเชิงสุ่มต่อเนื่อง}$$

ตัวอย่าง ก. จำนวนลูกค้าที่เข้ามาติดต่อธนาคารในคาบเวลา 09.00-10.00 มีการ**แจกแจงดังนี้**

x	90	100	110
f(x)	0.4	0.55	0.05

จงหาค่าคาดหวังของจำนวนลูกค้าที่เข้ามาติดต่อธนาคารในคาบเวลานั้น

$E(X) = 90(0.4) + 100(0.55) + 110(0.05)$

$= 36+55+5.5 = 96.5$

ข. กำหนดฟังก์ชันความหนาแน่นของตัวแปรเชิงสุ่ม X เป็นดังนี้

$f(x) = 2(x-1)\ ;\ 1 \leq x \leq 2$

สำหรับ E(X) หาได้ดังนี้

$$\int x f(x)dx = \int_1^2 x[2(x-1)]dx$$

$$= \int_1^2 (2x^2-2x)dx = (2x^3/3-x^2)\Big|_1^2$$

$$= (16/3-4)-(2/3-1) = 12/3$$

ค่าคาดหวังนั้นเราจะไม่แปลความหมายในทำนองว่าเกิดขึ้นบ่อย หรือน่าจะเป็นสูง แต่จะแปลว่าโดยเฉลี่ยแล้วค่าคาดหวังจะเป็นเท่านั้นเท่านี้ ซึ่งค่าคาดหวังนั้นจะเป็นค่าของตัวแปรเชิงสุ่มที่เกิดขึ้นบ่อยหรือไม่ก็ได้ และจะเป็นค่าของตัวแปรเชิงสุ่มหรือไม่ก็ได้ จากตัวอย่าง (1) เราจะเห็นว่าค่าคาดหวังเป็น 96.5 ซึ่งจะไม่เป็นค่าจริงของตัวแปรเชิงสุ่ม X (จำนวนลูกค้า) เลย

สำหรับค่าคาดหวังนั้นเรานำไปประยุกต์ในทางธุรกิจและเศรษฐศาสตร์เกี่ยวกับการลงทุนในโครงการหรือกิจการต่าง ๆ นั่นคือเราใช้เป็นเครื่องมืออย่างหนึ่งที่จะตัดสินว่าควรลงทุนประเภทไหนดีโดยปกติเราจะเลือกลงทุนในโครงการที่ให้ค่าคาดหวังสูงสุด

ค่าคาดหวังนั้นเราใช้เป็นมาตราวัดสรุปของตัวแปรเชิงสุ่มแบบอันตรภาคและอัตราส่วนเท่านั้น ไม่ใช่กับตัวแปรเชิงสุ่มนามบัญญัติและเรียงอันดับเป็นเด็ดขาด

ค่าคาดหวังมีชื่อเรียกต่าง ๆ กันดังนี้ Mathematical Expectation, Expectation, Ideal Average, Theoretical Mean, Mean Value, Average Value แต่ที่ใช้กันบ่อย ๆ คือ Mean และ Expected Value

อสมการของมาร์คอฟ (Markov's Inequality) บทบาทของค่าเฉลี่ยซึ่งเป็นพารามิเตอร์สำหรับวัดค่ากลาง ๆ ของการแจกแจงมีมากมาย โดยเฉพาะในทฤษฎีที่เรียกว่า อสมการของมาร์คอฟดังนี้

ทฤษฎี ถ้า X เป็นตัวแปรเชิงสุ่มที่มีค่าไม่เป็นลบ (Non-negative Values) และมีส่วนเฉลี่ย E(X) แล้ว

$$P(X \geq k) \leq E(X)/k$$

ในเมื่อ k เป็นจำนวนบวกใด ๆ $(k > 0)$

นั่นก็คือ $P(x \geq k) \leq E(X)/k \Leftrightarrow P\left\{X \geq kE(X)\right\} \leq 1/k$

หรือ $P\left\{X \leq k\,E(X)\right\} \geq 1 - 1/k \Leftrightarrow P(X \leq k) \geq 1 - E(X)/k$

สำหรับ $k \geq 1$ ทฤษฎีนี้จึงจะใช้ได้

ตัวอย่าง ฟังก์ชันน่าจะเป็นกำหนดไว้ดังนี้

x	1	2	3
f(X)	0.25	0.50	0.25

$$E(X) = 2$$

ดังนั้น $P(X \leq 2E(X))$ $P(X \leq 4) \geq 1-1/2 = 0.5$

แต่จากเป็นจริง $P(X \leq 4) = 1.0$

กฎเกี่ยวกับส่วนเฉลี่ย (Laws of Expected Value) มีกฎที่เป็นประโยชน์ในการคำนวณส่วนเฉลี่ยอยู่หลายกฎ กฎเหล่านี้จะช่วยให้เราคำนวณส่วนเฉลี่ยของสิ่งที่ต้องการในเทอมของส่วนเฉลี่ยที่เราทราบค่าแล้ว และกฎเหล่านี้จะใช้ได้ทั้งตัวแปรเชิงสุ่มทั้งสองประเภท

(1) ถ้า C เป็นค่าคงที่แล้ว $E(C) = C$

(2) ถ้า X เป็นตัวแปรเชิงสุ่ม และ C เป็นค่าคงที่แล้ว

$$E(X+C) = E(X)+C$$

(3) ถ้า X เป็นตัวแปรเชิงสุ่ม และ C เป็นค่าคงที่แล้ว

$$E(CX) = CE(X)$$

(4) ถ้า g(X) เป็นฟังก์ชันของตัวแปรเชิงสุ่ม X แล้ว

$$E(g(X)) = \begin{cases} \sum_{\forall x} g(x)\,f(x)\ ; X \text{ ไม่ต่อเนื่อง} \\ \int_{\forall x} g(x)\,f(x)dx\ ; X \text{ ต่อเนื่อง} \end{cases}$$

(5) ถ้า X และ Y เป็นตัวแปรเชิงสุ่มสองตัวแล้ว

$$E(X+Y) = E(X)+E(Y)$$

(6) ถ้า X และ Y เป็นตัวแปรเชิงสุ่มสองตัวที่เป็นอิสระกันแล้ว

$$E(XY) = E(X)\,E(Y)$$

ตัวอย่าง กำหนด X มีการแจกแจงดังนี้

x	0	1	2
f(x)	1/4	2/4	1/4

$E(X) = 0(1/4)+1(2/4)+2(1/4)$

สำหรับ (1) $E(X+5) = (0+5)\,1/4+(1+5)\,2/4+(2+5)\,1/4$

$= 5/4+12/4+7/4 = 24/4 = 6$

$= 1+5 = E(X)+5$

(2) $E(2X) = 2(0)(1/4)+2(1)(2/4)+2(2)(1/4)$

$= 0+4/4+4/4 = 8/4 = 2 = 2(1)$

$$= 2\,E(X)$$

(3) ถ้า $Y = X^2 + 3$

$$E(Y) = E(x^2 + 3) = (0^2 + 3)(1/4) + (1^2 + 3)(2/4) + (2^2 + 3)(1/4)$$

$$3/4 + 8/4 + 7/4 = 18/4 = 4.5$$

ข. **มาตราวัดตำแหน่ง** (Measures of Position) มาตราวัดตำแหน่งเป็นพารามิเตอร์อันดับ (order parameters) หรือค่าแบ่ง (Partition Value) ที่แบ่งการแจกแจงของตัวแปรเชิงสุ่มออกเป็นสัดส่วนต่าง ๆ มาตราวัดตำแหน่งที่เป็นรากของมาตราอื่นเรียกว่า ควอนไทล์ (Quantiles) หรือแฟรคไทล์ (Fractiles) ซึ่งกำหนดไว้ดังนี้

นิยาม: $\mathbf{x}_p$ จะเป็นควอนไทล์ที่ $p (0 < p < 1)$ ของตัวแปรเชิงสุ่ม X ถ้า

$$P(X \leq x_p) \geq p \text{ และ } P(X \geq x_p) \geq 1 - p$$

ถ้า X เป็นตัวแปรเชิงสุ่มต่อเนื่อง x_p จะหาได้จากสมการ

$$P(X \leq x_p) = p$$

สำหรับ (1) p ที่มีค่าเป็น 0.5 เราจะเรียก x_p ว่ามัธยฐานของ X

(2) p ที่มีค่าเป็น 0.25, 0.50 และ 0.75 เราจะเรียก x_p ว่าควอร์ไทล์ที่ 1, 2 และ 3 ของ x ตามลำดับ เราจะเรียกเป็น O Q_2 และ Q_3 แต่เรามักเรียก Q_1 และ Q_3 ว่าควอร์ไทล์ต่ำ (Lower Quartile) และควอร์ไทล์สูง (Upper Quartile) ตามลำดับ

(3) p ที่มีค่าเป็น 0.1, 0.2, 0.3, ..., 0.9 เราเรียก x_p ว่าเดไซล์ (Decile) ที่ 1, 2, ..., 9 ของ X ตามลำดับ เราจะเขียน $D_1, D_2, \ldots D_9$ แทน

(4) p ที่มีค่าเป็น 0.01, 0.02, 0.03, ..., 0.99 เราจะเรียก x_p ว่าเปอร์เซ็นไทล์ (Percentile) ที่ 1, 2, 3, ..., 99 ของ X ตามลำดับ เราจะเรียกว่า $P_1, P_2, \ldots, P_{99}$ ตามลำดับ

มาตราวัดตำแหน่งนี้ใช้สรุปตัวแปรเชิงสุ่มแบบเรียงอันดับ, อันตรภาค และอัตราส่วน ไม่ใช้กับตัวแปรเชิงสุ่มแบบนามบัญญัติเด็ดขาด

ตัวอย่าง กำหนดตัวแปรเชิงสุ่ม X มีการแจกแจงดังนี้

$$f(x) = 1/\beta \;;\; \alpha < x < \alpha + \beta$$

จงหาควอนไทล์ที่ p เมื่อ p = 0.2, 0.5 และ 0.75

ให้ x_p เป็นควอนไทล์ที่ p

เมื่อ (1) $p = 0.2$ เราได้ $\int_{\alpha}^{x_{0.2}} (1/\beta)\,dx = 0.2$

$$(x/\beta)\Big|_{\alpha}^{x_{0.2}} = 0.2$$

$$(x_{0.2}/\beta) - \alpha/\beta = 0.2$$

$$x_{0.2} = 0.2\,\beta + \alpha$$

(2) $p = 0.5$ เราจะได้ $x_{0.5} = 0.5\beta + \alpha$

และ (3) $p = 0.75$ เราจะได้ $x_{0.75} = 0.75\beta + \alpha$

3.5.2 **มาตราวัดการกระจาย** (Measure of Dispersion or Concentration) มาตราวัดการกระจายเป็นปริมาณหรือพารามิเตอร์ที่แสดงการกระจัดกระจาย (spread or scatter) หรือแสดงจำนวนความแตกต่างระหว่างค่าที่เป็นไปได้ของตัวแปรเชิงสุ่ม มาตราวัดการกระจายที่น่าจะพูดถึงมีดังนี้

(1) พิสัย (Range) เป็นมาตราวัดการกระจายที่หยาบแต่คำนวณได้รวดเร็ว และใช้สรุปตัวแปรเชิงสุ่มแบบอันตรภาคและอัตราส่วนเท่านั้น พิสัยกำหนดไว้ว่า

นิยาม: ถ้า X เป็นตัวแปรเชิงสุ่มแล้ว พิสัย R จะเป็นผลต่างของค่าของตัวแปรเชิงสุ่มที่มากที่สุดกับน้อยที่สุด นั่นคือ

$$R = Max(x_i) - Min(x_i)$$

ตัวอย่าง $f(x) = \frac{1}{\beta - \alpha} \quad \alpha < x < \beta$

ดังนั้น $R = \beta - \alpha$ เพราะ $Max(x_i) = \beta$ และ $Min(x_i) = \alpha$

(2) ส่วนเบี่ยงเบนควอร์ไทล์ (Quartile Deviation) เป็นระยะทางครึ่งหนึ่งของระยะทางระหว่างควอร์ไทล์ที่ 1 กับ ควอร์ไทล์ที่ 3 นั่นคือ

$$QD = (Q_3 - Q_1)/2$$

บางทีเราเรียก QD ว่า กึ่งพิสัยควอร์ไทล์ (Semi-Inter-quartile Range) เพราะ $Q_3 - Q_1$ เป็นพิสัยควอร์ไทล์ (Quartile Range)

ส่วนเบี่ยงเบนควอร์ไทล์ QD นี้ใช้เป็นมาตราวัดการกระจายของตัวแปรเชิงสุ่มแบบเรียงอันดับ อันตรภาค หรือ อัตราส่วน

ตัวอย่าง กำหนด $f(x) = 1/\beta$; $\alpha < x < \alpha + \beta$

$Q_1 = x_{0.25} = 0.25\beta + \alpha$

$Q_3 = x_{0.75} = 0.75\beta + \alpha$

ดังนั้น $QD = \frac{1}{2}(0.75\beta - 0.25\beta) = \frac{1}{2}(0.5\beta)$

$= 0.25\beta$

มาตราวัดการกระจายที่มีลักษณะเดียวกับ QD ก็คือส่วนเบี่ยงเบนเดไซล์ (Decile Deviation, DD) ซึ่งกำหนดไว้ว่า

$$DD = (D_9 - D_1)/2 = (P_{90} - P_{10})/2$$

(3) ส่วนเบี่ยงเบนเฉลี่ย (Mean Deviation, MD.) ส่วนเบี่ยงเบนเฉลี่ยเป็นค่าเฉลี่ยสมบูรณ์ (absolute) ของส่วนเบี่ยงเบน (Deviation) ที่ค่าของตัวแปรเชิงสุ่มห่างจากค่าเฉลี่ยของมัน P.S. Laplace (1749 - 1827) เป็นผู้เสนอเทคนิคของมาตราวัดแบบนี้ MD กำหนดไว้ว่า

นิยาม: ถ้า X เป็นตัวแปรเชิงสุ่มซึ่งมีค่าเฉลี่ย $E(X)=\mu$ แล้ว

$$MD = E\left|X-\mu\right| = \begin{cases} \sum\limits_{\forall x} \left|x-\mu\right| f(x) \text{ ; ไม่ต่อเนื่อง} \\ \int_x \left|x-\mu\right| f(x)dx \text{ ; ต่อเนื่อง} \end{cases}$$

เนื่องจาก MD เกี่ยวข้องกับค่าสมบูรณ์ (absolute value) ของส่วนเบี่ยงเบนในทางคณิตศาสตร์ถือว่ายุ่งยากในทางคำนวณ จึงต้องหลีกเลี่ยงไปใช้วิธีอื่นคือ ความแปรปรวนหรือส่วนเบี่ยงเบนมาตรฐานซึ่งมีความสำคัญจะได้กล่าวต่อไป

ส่วนเบี่ยงเบนเฉลี่ยนั้นไม่ใช้เป็นมาตราวัดการกระจายของตัวแปรเชิงสุ่มแบบนามบัญญัติหรือเรียงอันดับ

ตัวอย่าง กำหนดจำนวนสินค้าที่ลูกค้าแต่ละรายสั่งซึ่งเป็นตัวแบบเชิงสุ่มมีการแจกแจงดังนี้

x	10	20	30
f(x)	0.5	0.3	0.2

สำหรับ MD. หาได้ดังนี้

$$E(X) = 10(0.5)+20(0.3)+30(0.2)=17$$

$$MD. = E\left|X-17\right|$$

$$= |10-17|\ (0.5)+|20-17|\ (0.3)+|30-17|\ (0.2)$$

$$= 3.5+0.9+2.6=7$$

(4) ความแปรปรวน (Variance) ความแปรปรวนเป็นมาตราวัดการกระจายที่มนุษย์รู้จักกันมากและใช้กันมาก K.F. Gauss (1777 - 1855) เป็นผู้พัฒนาคนแรกจากส่วนเบี่ยงเบนโดยการเฉลี่ย กำลังสองของส่วนเบี่ยงเบนจากค่าเฉลี่ย

นิยาม: ถ้า X เป็นตัวแปรเชิงสุ่มที่มี $E(X)=\mu$ แล้ว ความแปรปรวน, $V(X)=\sigma^2$ กำหนดไว้ดังนี้

$$V(X)=E(X-\mu)^2 = \begin{cases} \sum\limits_{\forall x}(x-\mu)^2 f(x) \text{ ; ไม่ต่อเนื่อง} \\ \int_{\forall x}(x-\mu)^2 f(x)dx \text{ ; X ต่อเนื่อง} \end{cases}$$

ตัวอย่าง กำหนด X เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงดังนี้

x	0	1	2
f(x)	0.25	0.50	0.25

$$E(X) = 0(0.25)+1(0.50)+2(0.25)=1$$

$$V(X) = (0-1)^2(0.25)+(1-1)^2(0.50)+(2-1)^2(0.25)$$

$$= 0.25+0+0.25=0.50$$

จากนิยามของความแปรปรวนนี้ ถ้า E(X) มีจุดทศนิยมจะทำให้การคำนวณไม่สะดวก เราจึงมักใช้สูตรต่อไปนี้แทน

$$V(X) = E(X)^2 - \mu^2$$

สำหรับตัวอย่างข้างบนนี้เราจะหาความแปรปรวนด้วยสูตรนี้ จะได้เป็น

$$V(X) = (0^2(0.25) + 1^2(0.50) + 2^2(0.25)) - (1)^2$$

$$= (0+0.50+1) -1 = 1.50-1 = 0.50$$

คุณสมบัติของความแปรปรวน (Some Properties of Variance) ถ้ากำหนดให้ C เป็นค่าคงที่ และ X เป็นตัวแปรเชิงสุ่มใด ๆ โดยที่ $V(X) \geq 0$ กฎต่าง ๆ ต่อไปนี้จะเป็นประโยชน์ต่อการคำนวณความแปรปรวนฟังก์ชั่นของตัวแปรเชิงสุ่ม

(1) ถ้า g(X) เป็นฟังก์ชั่นของ X แล้ว V(g(X)) กำหนดไว้ดังนี้

$$V(g(X)) = E[g(X) - E\, g(X)]^2$$

(2) $V(C) = 0$

(3) $V(X+C) = V(X)$

(4) $V(CX) = C^2 V(X)$

(5) ถ้า X และ Y เป็นตัวแปรเชิงสุ่มที่เป็นอิสระกันแล้ว

$$V(X+Y) = V(X) + V(Y) \quad \text{หรือ} \quad V(X-Y) = V(X) + V(Y)$$

(6) $V(X) \leq E(X-C)^2$

ในทางธุรกิจและเศรษฐศาสตร์จะใช้การกระจัดกระจาย (Spread or variability) เป็นมาตราวัดการเสี่ยง (Risk) คือ ถ้ากระจัดกระจายมากก็เสี่ยงมาก โดยทั่วไปเราใช้ความแปรปรวนวัดการเสี่ยง ในแง่ของการลงทุนที่เรากล่าวมาแล้วว่า เราพิจารณาจากค่าคาดหวังของผลตอบแทนทางหนึ่ง เมื่อถึงตอนนี้เราก็มีเครื่องมือช่วยอีกอย่างหนึ่งคือความแปรปรวนหรือการเสี่ยง นั่นคือโครงการไหนเสี่ยงมากเราก็ไม่ลงทุนในโครงการนั้น ถ้าในกรณีที่เราพบโครงการที่มีค่าคาดหวังของผลตอบแทนมากแต่มีการเสี่ยงมากเราก็พิจารณายาก ดังที่ H. Markowitz กล่าวไว้ว่า "ระหว่างการลงทุน X กับ Y ถ้ามี $E(Y) > E(X)$ และการเสี่ยงของ Y มากกว่าของ X หรือ $V(Y) > V(X)$ แล้วเป็นการยากมากที่จะเลือกระหว่างการลงทุนทำนองนั้น"

ความแปรปรวนใช้เป็นมาตราวัดสรุปเกี่ยวกับการกระจายของตัวแปรเชิงสุ่มแบบอันตรภาคและอัตราส่วนเท่านั้น

(5) ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) ความยุ่งยากในการใช้ความแปรปรวนมีอยู่ประการหนึ่งคือ มันไม่ได้วัดการกระจายในหน่วย (Unit) เดียวกับหน่วยของค่าของตัวแปรเชิงสุ่ม มิติ (Dimensionality) ของความแปรปรวนจะเป็นกำลังสองของมิติของตัวแปรเชิงสุ่มของมันเอง เช่น ถ้า X เป็นนิ้ว E(X) จะเป็นนิ้ว แต่ความแปรปรวนวัดเป็นนิ้วกำลังสองเพื่อให้มาตราวัดการ

กระจายมีหน่วยเช่นเดียวกับของค่าของตัวแปรเชิงสุ่ม เราจึงกำหนดมาตราวัดอันหนึ่งซึ่งเป็นรากที่สองของความแปรปรวน เราจะเรียกมันว่าส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) ของตัวแปรเชิงสุ่มซึ่งนิยามไว้ว่า

นิยาม: ถ้า X เป็นตัวแปรเชิงสุ่มที่มีความแปรปรวน V(X) แล้วส่วนเบี่ยงเบนมาตรฐานกำหนดไว้ดังนี้

$$\sigma = \sqrt{V(X)} = \sqrt{E(X^2) - (E(X))^2}$$

ตัวอย่าง กำหนดฟังก์ชันสำหรับ X เป็นดังนี้

$$f(x) = 2x/3, \quad 1 \leq x \leq 2$$

เราจะได้ $E(X) = \int_1^2 x(2x/3)dx = 2x^3/9 \Big|_1^2 = 14/9$

$$E(X^2) = \int_1^2 x^2(2x/3)dx = x^4/6 \Big|_1^2 = 15/6$$

ดังนั้น $V(X) = E(X^2) - (E(X))^2 = 15/6 - (14/9)^2$

$$= 13/162$$

และ $\sigma = \sqrt{13/162} = 0.283$

อสมการของเชบบีเชฟ (Tchebycheff's Inequality) บทบาทของส่วนเบี่ยงเบนมาตรฐานในฐานะที่เป็นมาตรวัดการกระจายดูเหมือนจะมีความสำคัญมากเมื่อ P.L. Tchebycheff (1881 – 1894) นักคณิตศาสตร์ชาวรัสเซีย ได้พัฒนาทฤษฎีอันหนึ่งโดยการใช้ส่วนเบี่ยงเบนมาตรฐาน ทฤษฎีนี้ชื่อว่า Tchebycheff's Inequality และกล่าวไว้ดังนี้

ทฤษฎี ความน่าจะเป็นที่ค่าของตัวแปรเชิงสุ่ม X จะเบี่ยงเบนจากค่าเฉลี่ยของมันมากกว่า k ส่วนเบี่ยงเบนมาตรฐานจะเท่ากับหรือน้อยกว่าส่วนกลับของ k ส่วนเบี่ยงเบนมาตรฐานจะเท่ากับหรือมากกว่า $1 - 1/k^2$

นั่นคือ $P(|X - E(X)| \leq k\sigma) \geq 1 - 1/k^2$

ถ้า k อยู่ระหว่าง 0 กับ 1 จะไม่ให้อะไร เพราะจะได้ความน่าจะเป็นมีค่าเป็นลบ ลองพิจารณาค่าของ k ต่าง ๆ กันดังนี้

ความน่าจะเป็นจาก อสมการเชบบิเชฟ

k	ช่วง	ความน่าจะเป็น
1.25	$\mu \pm 1.25\sigma$	0.36
1.5	$\mu \pm 1.50\sigma$	0.56
2	$\mu \pm 2\sigma$	0.25
2.5	$\mu \pm 2.5\sigma$	0.84

3	$\mu \pm 3\sigma$	0.89
4	$\mu \pm 4\sigma$	0.94

แต่ทฤษฎีนี้จะให้ความน่าจะเป็นที่ X จะอยู่ในช่วง $\mu \pm k\sigma$ หรือช่วงกว้าง $2k\sigma$ (ซึ่งต้องรวมจุดกลาง E(X) ไว้ด้วย) เท่านั้น Frechet จึงได้ปรับปรุงทฤษฎีเป็นรูปทั่ว ๆ ไป เพื่อหาความน่าจะเป็นในช่วงอื่น ๆ ซึ่งไม่จำเป็นต้องรวมชุดกลางหรือ E(X) ไว้ด้วยดังต่อไปนี้

ทฤษฎี ถ้าตัวแปรเชิงสุ่ม X มีค่าเฉลี่ย μ และส่วนเบี่ยงเบนมาตรฐานแล้วความน่าจะเป็นที่ X จะมีค่าอยู่ในช่วง $\mu - k_1\sigma$ และ $\mu + k_2\sigma$ จะเท่ากันหรือมากกว่า

$$1 - \frac{1 + ((k_2-k_1)/2)^2}{((k_2+k_1)/2)^2} \quad ; \quad k_1, k_2 > 0$$

นั่นก็คือ $P(\mu - k_1\sigma < X < \mu + k_2\sigma) \geq 1 - \frac{1 + ((k_2-k_1)/2)^2}{((k_2+k_1)/2)^2}$

ตัวอย่าง ให้ X เป็นตัวแปรเชิงสุ่มที่แทนจำนวนขนมปังที่ขายได้แต่ละวัน ซึ่งมี E(X) = 25 ปอนด์ และ σ = 5 ปอนด์

อสมการเชบบีเชฟ จะกล่าวว่าความน่าจะเป็นที่จะขายขนมปังได้ระหว่าง 25 ถึง 35 ปอนด์ จะเท่ากับ 75% หรือมากกว่า

แต่จากการทดลองหรือในการปฏิบัติเราได้กฎการทดลองซึ่งจะให้ความน่าจะเป็นหรือโอกาสของตัวแปรเชิงสุ่มในช่วงต่าง ๆ ถ้าตัวแปรเชิงสุ่มมีการแจกแจงแบบระฆัง ดังนี้

กฎการทดลอง (Empirical Rule) ถ้าการแจกแจงของตัวแปรเชิงสุ่มซึ่งมีรูปร่างคล้ายระฆัง (Bell-Shaped distribution) แล้วโอกาสที่ค่าของตัวแปรเชิงสุ่มจะอยู่ในช่วง $\mu + \sigma$, $\mu + 2\sigma$ และ $\mu + 3\sigma$ ประมาณ 68%, 95%, 99.7% ตามลำดับ

(6) สัมประสิทธิ์ของการกระจาย (Coefficient of Dispersion) สัมประสิทธิ์ของการกระจายเป็นมาตราวัดที่ใช้เปรียบเทียบการกระจายของตัวแปรเชิงสุ่มสองตัว มันเป็นจำนวนเลขแท้ (Pure number) ที่ไม่มีหน่วยนั่นเอง จึงสามารถเปรียบเทียบตัวแปรเชิงสุ่มที่มีขนาดของค่าตัวแปรโตต่างกัน หรือมีหน่วยของค่าตัวแปรเชิงสุ่มต่างกันได้ เช่นค่าของตัวแปรเชิงสุ่มตัวหนึ่งมีหน่วย 1,000 บาท อีกตัวหนึ่งมีหน่วย 10 บาท ก็สามารถเปรียบเทียบการกระจายกันได้ หรือค่าของตัวแปรเชิงสุ่มตัวหนึ่งมีหน่วยเป็นปอนด์ อีกตัวหนึ่งมีหน่วยเป็นบาท เป็นต้น สัมประสิทธิ์ของการกระจายกำหนดไว้ดังนี้

$$\text{สัมประสิทธิ์ของการกระจาย} = \frac{\text{มาตราวัดการกระจาย}}{\text{มาตราวัดค่ากลางที่เหมาะสม}}$$

สัมประสิทธิ์ของการกระจายที่สำคัญดังนี้

(ก) สัมประสิทธิ์ของความผันแปร (Coefficient of Variation CV.) มาตราวัดนี้ Karl Pearson (1957 - 1936) เป็นผู้พัฒนาขึ้นมาดังสูตรต่อไปนี้

$$CV = \frac{\sigma}{\mu}(100)$$

ถ้า $\mu \rightarrow 0$ มาตราวัดนี้จะใช้ได้ไม่ดี

ตัวอย่าง รายได้จากการขายสินค้าของบริษัทใหญ่และเด็กมีส่วนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานดังนี้

	บริษัทใหญ่	บริษัทเล็ก
ส่วนเฉลี่ย	2,000,000	100,000
ส่วนเบี่ยงเบนมาตรฐาน	200,000	10,000

ถ้ามองเผิน ๆ จะเห็นว่าการกระจายของรายได้ในบริษัทใหญ่จะมากกว่า แต่ที่จริงแล้วทั้งสองมีการกระจายเท่า ๆ กันคือ

บริษัทใหญ่ $CV = (200,000/2,000,000)(100) = 10\ \%$

บริษัทเล็ก $CV = (10,000/100,000)(100) = 10\ \%$

(ข) สัมประสิทธิ์ของส่วนเบี่ยงเบนควอร์ไทล์ VQ (Coefficient of Quartile Deviation) มาตราวัด VQ นี้กำหนดไว้ว่า

$$VQ = \frac{Q_3 - Q_1}{Q_3 + Q_1}(100)$$

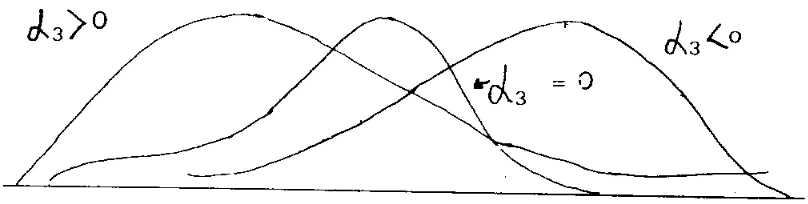
(ค) สัมประสิทธิ์ที่ขึ้นอยู่กับส่วนเบี่ยงเบนเฉลี่ย (Vm) สำหรับ Vm กำหนดไว้ดังนี้

$$\begin{aligned} Vm &= (MD/\mu)(100) \\ &= (MD/Median)(100) \end{aligned}$$

3.5.3 มาตราวัดความเบ้และความสูง (Measures of Skewness and Kurtosis) ความเบ้นี้ใช้พิจารณาความไม่สมมาตรหรือความเบ้ของการแจกแจง R.A. Fisher ได้เสนอสูตรไว้ดังนี้

$$\alpha_3 = \mu_3/\sigma^3 \ ; \quad \mu_3 = E(X - \mu)^3$$

ถ้าการแจกแจงที่มีมัธยฐานเดียว และ $\alpha_3 = 0$ แล้วการแจกแจงจะสมมาตรถ้า $\alpha_3 > 0$ การแจกแจงของตัวแปรเชิงสุ่มจะเบ้ทางบวกหรือขวา (Positive Skewness) แต่ถ้า $\alpha_3 < 0$ จะเบ้ทางซ้ายหรือลบ (Negative Skewness)



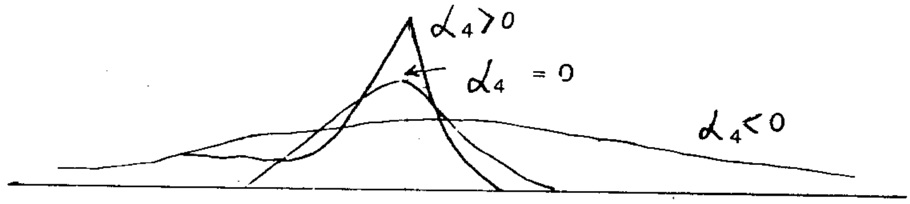
Karl Pearson เป็นคนแรกที่ได้เสนอมาตราวัดความเบ้ ไว้ดังนี้

$$S_K = (\mu - X_0)/\sigma ; X_0 : \text{มัธยฐาน}$$

มาตราวัดความสูงเป็นดัชนีที่บอกให้ทราบว่าแจกแจงของตัวเชิงสุ่มมีแนวโน้มจะสูงกว่าหรือแบน (Peaked or Flat) กว่าการแจกแจงปกติ มาตราวัดความสูงนิยามไว้ดังนี้

$$\alpha_4 = \mu_4/\sigma^4 - 3 \; ; \; \mu_4 = E(X - E(X))^4$$

ถ้า $\alpha_4 > 0$ จะแสดงว่าการแจกแจงจะสูงโด่งกว่าการแจกแจงปกติ แต่ถ้าเป็นลบจะแบนกว่าดังรูป



เมื่อ $\alpha_4 = 0, \alpha_4 > 0$, หรือ $\alpha_4 < 0$ การแจกแจงนั้นจะเรียกว่า Mesokurtic, Leptokurtic หรือ Platykurtic ตามลำดับ

3.6 ตัวแปรเชิงสุ่มมาตรฐาน (Standardized Random Variable)

บางครั้งเราต้องการเปรียบเทียบค่าของตัวแปรเชิงสุ่มคนละตัวกัน หรือที่มีหน่วยของค่าตัวแปรเชิงสุ่มต่างกัน วิธีการเปรียบเทียบเราทำได้โดยแปลงค่าตัวแปรเชิงสุ่มให้เป็นมาตรฐานก่อน นิยามของตัวแปรเชิงสุ่มมาตรฐานกำหนดไว้ดังนี้

นิยาม ถ้า X เป็นตัวแปรเชิงสุ่มที่มีค่าเฉลี่ย μ และส่วนเบี่ยงมาตรฐาน σ แล้วตัวแปรเชิงสุ่ม Z ซึ่งกำหนดว่า $Z = (X - \mu)/\sigma$ จะเป็นตัวแปรเชิงสุ่มมาตรฐาน และตัวแปรเชิงสุ่มมาตรฐานจะต้องมีค่าเฉลี่ยเป็น 0 และความแปรปรวนเป็น 1 นั่นคือ $E(Z) = 0$ และ $V(Z) = 1$

3.7 โมเมนต์และฟังก์ชันกำเนิดของตัวแปรเชิงสุ่ม (Moments and Generating Functions)

ค่าคาดหวังของกำลังของตัวแปรเชิงสุ่มที่มีการแจกแจงจะเรียกว่า โมเมนต์ของการแจกแจง โมเมนต์ต่าง ๆ ของตัวแปรเชิงสุ่มจะอธิบายคุณลักษณะของฟังก์ชันแจกแจงน่าจะเป็นได้ดี **เพราะ** โมเมนต์ของฟังก์ชันการแจกแจงหนึ่ง ๆ จะมีเพียงตัวเดียว ถ้าตัวแปรเชิงสุ่มสองตัวมีโมเมนต์**เดียวกัน** แล้วมักจะมีฟังก์ชันแจกแจงเหมือนกัน หรือในทางกลับกัน และโมเมนต์ของตัวแปร**เชิงสุ่มยังจะ**ให้เครื่องมือทางทฤษฎีที่สำคัญอีก นิยามของโมเมนต์ของตัวแปรเชิงสุ่ม**เป็นดังนี้**

นิยาม ค่าคาดหวังของฟังก์ชัน $g(X) = X^k$ หรือ m_k

$m_k = E(X^k), k = 0, 1, 2, 3, ...$

จะเรียกว่าโมเมนต์ลำดับที่ k ของตัวแปรเชิงสุ่ม X

สำหรับค่าคาดหวังและความแปรปรวนที่กล่าวมาแล้วก็เป็นโมเมนต์อย่างหนึ่งคือ เมื่อ $k = 1$, แล้ว $m_1 = E(X)$ และ $m_2 = E(X)^2$

ดังนั้น $V(X) = E(X)^2 - (E(X))^2 = m_2 - (m_1)^2$

หรือ $m_2 = V(X) + (E(X))^2$

ตัวอย่าง (1) กำหนดฟังก์ชั่นน่าจะเป็นดังนี้

$$p(X) = 0.1, x = 0$$
$$= 0.9, x = 1$$
$$m_k = E(X^k) = m^k o^k (0.1) + 1^k (0.9)$$
$$= 0.9, k = 1, 2, 3, ...$$

นั่นคือ ตัวแปรเชิงสุ่ม X มีโมเมนต์คงที่เป็น 0.9 สำหรับทุกค่า k

(2) ให้ X เป็นเวลาที่ลูกค้าใช้ในการติดต่อธนาคารมีการแจกแจงดังนี้

$$f(x) = 1/10 ; 20 < x < 30$$

$$m_k = E(X^k) = \int_{20}^{30} x^k (1/10) dx$$

$$= \frac{30^{k+1} - 20^{k+1}}{10(k+1)}, k = 1, 2, 3, ...$$

ดังนั้น $m_1 = \frac{(30)^2 - (20)^2}{10(2)} = 25$

$m_2 = \frac{(30)^3 - (20)^3}{10(3)} = 633\frac{1}{3}$

นั่นคือ $E(X) = 25$ และ $\sigma^2 = 8\frac{1}{3}$

โมเมนต์ที่กล่าวมานี้บางทีเรียกว่า โมเมนต์รอบจุดกำเนิด หรือโมเมนต์ธรรมดา (Ordinary moments) นอกจากนี้ยังมีโมเมนต์ต่าง ๆ อีก ดังนิยามต่อไปนี้

นิยาม โมเมนต์ลำดับที่ k รอบจุด C ใด ๆ กำหนดไว้ว่า

$M_k = E(X - C)^k ; k=1,2,...$ **ถ้า $C = \mu$ แล้ว**

M_k เป็นโมเมนต์ที่มีความสำคัญมากซึ่งนิยามไว้ดังนี้

นิยาม โมเมนต์ลำดับที่ k รอบค่าเฉลี่ย จะเรียกว่าโมเมนต์กลาง (Central Moment) กำหนดไว้ดังนี้

สำหรับ $k = 1, 2, 3, 4$ เราได้ $M_1 = 0$, $M_2 = V(X)$ μ_3 = ความเบ้ (skewness) ซึ่งใช้เป็นมาตราวัดการไม่สมมาตร (Asymmetry) ของการแจกแจง และ μ_4 = ความสูง (Kurtosis) ซึ่งใช้เป็นมาตราวัด

ความสูง(Peak or Flat) ของการแจกแจง

โดยการใช้ทฤษฎีทวินาม (Bionomial Theorm) เราสามารถหาความสัมพันธ์ระหว่างโมเมนต์ของตัวแปรเชิงสุ่มกับโมเมนต์รอบค่าเฉลี่ยของตัวแปรเชิงสุ่มได้ดังนี้

$$\begin{aligned}\mu_k &= E(X-\mu)^k \\ &= E\sum_{i=0}^{k}\binom{k}{i}X^i(-\mu)^{k-i} \\ &= \sum_{i=0}^{k}\binom{k}{i}m_i(-\mu)^{k-i}\end{aligned}$$

ดังนั้น
$$\begin{aligned}\mu_1 &= m_1-\mu = 0 ;\ m_1 = \mu \\ \mu_2 &= (-\mu)^2 + 2m_1(-\mu) + m_2 \\ &= m_2-\mu^2 \\ \mu_3 &= (-\mu)^3 + 3m_1(-\mu)^2 + 3m_2(-\mu) + m_3 \\ &= m_3 - 3m_2\mu + 2\mu^3\end{aligned}$$

ในทำนองเดียวกันเราจะได้

$$\begin{aligned}m_k &= E[X^k] = E[(X-\mu)+\mu]^k \\ &= E\sum_{i=0}^{k}\binom{k}{i}(x-\mu)^i\mu^{k-i} \\ &= E\sum_{i=0}^{k}\binom{k}{i}\mu_i\mu^{k-i}\end{aligned}$$

ดังนั้น
$$\begin{aligned}m_1 &= \mu + \mu_1 = \mu ;\ \mu_1 = 0 \\ m_2 &= m_1^2 + 2\mu_1\mu + \mu_2 \\ &= \mu^2 + \mu_2 \\ m_3 &= \mu^3 + 3\mu_1\mu^2 + 3\mu_2\mu + \mu_3 \\ &= \mu_3 + \mu_2\mu + \mu_3\end{aligned}$$

นอกจากโมเมนต์ดังกล่าวก็ยังมีโมเมนต์ต่าง ๆ อีกดังนี้

(1) โมเมนต์ผลคูณ (Factorial moment) ซึ่งกำหนดไว้ดังนี้

$$\mu'_k = E(X^{(k)}) = E[X(X-1)\ldots\ldots(X-k+1)]$$

เมื่อ $k = 1, 2$ เราได้ $\mu'_1 = E(X)$

$$\mu'_2 = E[X(X-1)] = E(X)^2 - E(X) = m_2 - m_1$$

(2) โมเมนต์สมบูรณ์ (Absolute Moments) กำหนดไว้ดังนี้

$$\mu_k = E|X|^k$$

และโมเมนต์สมบูรณ์รอบค่าเฉลี่ย (Absolute Central Moment) คือ $\nu_k = E|X-\mu|^k$

เมื่อ $k = 1 = \nu_1 = M.D$

มีฟังก์ชันพิเศษบางอย่างใช้ผลิต (Generate) โมเมนต์ของตัวแปรเชิงสุ่มได้โดยการหา

อนุพันธ์ (Differentiation) ฟังก์ชั่นโมเมนต์มีดังนี้

- ฟังก์ชั่นกำเนิดโมเมนต์
- ฟังก์ชั่นกำเนิดแฟคเตอร์ของโมเมนต์
- ฟังก์ชั่นแสดงคุณลักษณะ

(1) ฟังก์ชั่นกำเนิดโมเมนต์ (Moment - Generating Function, M(t)) กำหนดไว้ดังนี้

นิยาม ฟังก์ชั่นกำเนิดโมเมนต์สำหรับตัวแปรสุ่ม X ที่มีค่าเฉลี่ย μ จะเป็น

$$M(t) = E[e^{tx}]$$

ในเมื่อ t เป็นเลขจำนวนจริงอยู่ในช่วงหนึ่งที่ทำให้ M(t) มีจริง (exist) ถ้าหา M(t) ได้แล้ว M(t) จะเป็นฟังก์ชั่นแบบต่อเนื่องที่สามารถหาอนุพันธ์ได้ เมื่อเราหาอนุพันธ์ต่าง ๆ โดยเทียบกับ t และคำนวณที่ t = 0 เราจะได้โมเมนต์ต่าง ๆ ดังนี้

$$M^k(t) = \left.\frac{d^k M(t)}{dt^k}\right|_{t=0}; \quad k = 1,2,3\ldots$$

$$M^1(t) = \frac{d}{dt}E[e^{tx}] = E[xe^{tx}]$$

ดังนั้น $M^1(0) = E(x) = m_1,\ M^2(t) = E(x^2e^{tx})$

$\ldots M^k(t) = E(x^ke^{tx}),\ M^k(0) = E(x^k) = m_k$

เราจะเห็นได้ว่าฟังก์ชั่นกำเนิดของ X จะผลิตโมเมนต์ของ X

ตัวอย่าง จำนวนลูกค้าที่เข้ามารับบริการในธนาคารสำหรับช่วงเวลาหนึ่งเป็นดังนี้

$$f(x) = e^{-a}a^x/x!\ ;\ x = 0, 1, 2, 3, \ldots\ldots$$

ดังนั้น $M(t) = E(e^{tx}) = \sum_{x=0}^{\infty} e^{xt}e^{-a}a^x/x!$

$$= e^{-a}\sum_{x=0}^{\infty}(ae^t)^x/x! = e^{-a}e^{ae^t}$$

$$M^1(0) = a = m_1$$

$$M^2(0) = a(1+a) = m_2$$

นอกจากฟังก์ชั่นกำเนิดโมเมนต์ของ X จะผลิตโมเมนต์รอบจุดกำเนิดได้ เรายังจะสามารถผลิตโมเมนต์รอบจุดอื่น ๆ ได้อีก ดังต่อไปนี้

กำหนด $M_{x-c}(t) = E[e^{t(x-c)}]$

$$= e^{-tc}E(e^{tx})$$

ดังนั้น $M_{x-c}(t) = e^{-tc}M_x(t)$

นั่นคือถ้าเราทราบ M(t) ของตัวแปรเชิงสุ่ม เราก็เอา e^{-tc} คูณเข้าไป ถ้า $c = \mu$ เราจะได้โมเมนต์รอบค่าเฉลี่ย

ตัวอย่าง กำหนด $f(y) = 100e^{-100y}\ ;\ y > 0$

$$M_y(t) = \frac{1000}{1000-t}; t < 1000$$
$$E(Y) = 1/1000$$

เมื่อเราต้องการหา $M_z(t)$ เมื่อ $Z = Y - 1/1000$

เราจะได้ $M_z(t) = e^{-t/1000} \left(\frac{1000}{1000-t}\right)$

ซึ่ง จะผลิตโมเมนต์รอบค่าเฉลี่ยได้ดังนี้

$$M_Z^1(0) = 0, \; M_Z^2(0) = \frac{1}{(1000)^2}$$

ที่เรากล่าวว่า $M(t)$ จะแสดงคุณลักษณะของฟังก์ชั่นน่าจะเป็นของตัวแปรเชิงสุ่มนั้น เราจะดูได้จากทฤษฎีต่อไปนี้

ทฤษฎี ให้ X และ Y เป็นตัวแปรเชิงสุ่มที่มีฟังก์ชั่นน่าจะเป็น $f(X)$ และ $g(X)$ ตามลำดับ สมมุติว่า $M(t)$ ของทั้ง X และ Y หาได้ (exist) และเท่ากันสำหรับทุกค่าของ t ที่อยู่ในช่วงหนึ่ง แล้วฟังก์ชั่นน่าจะเป็นของทั้งสองจะเท่ากัน (ยกเว้นบางจุดที่ไม่ต่อเนื่อง ถ้า X และ Y เป็นตัวแปรเชิงสุ่มต่อเนื่อง)

นั่นคือ $f(x) = f(y) \iff M_x(t) = M_y(t)$

ตัวอย่าง ให้ X เป็นตัวแปรเชิงสุ่มที่มี $M_x(t) = 1/(1-t)$ ซึ่ง $M(t)$ จะหาได้เมื่อ $-1 < t < 1$ แต่ Y มีฟังก์ชั่นหนาแน่นดังนี้ $g(Y) = e^{-Y}, y > o$ และมี $M_y(t)$ ดังนี้

$$M_y(t) = \int_0^\infty e^{yt} e^{-y} dy = 1/(1-t)$$

ดังนั้นตามทฤษฎีเราจะได้ว่า X จะต้องมีฟังก์ชั่นหนาแน่นดังนี้

$$f(x) = e^{-x}, x > 0$$

(2) ฟังก์ชั่นกำเนิดแฟคเตอร์ของโมเมนต์ (Generating function, G(t)) เป็นฟังก์ชั่นที่จะผลิตโมเมนต์ ผลคูณของตัวแปรเชิงสุ่มสำหรับตัวแปรเชิงสุ่มแบบไม่ต่อเนื่อง ฟังก์ชั่นนี้จะช่วยให้หาโมเมนต์ได้ง่าย เราจะนิยาม G(t) ของตัวแปรเชิงสุ่มได้ดังนี้

นิยาม สำหรับทุกค่าของ t ในช่วงจำกัดที่แน่นอน (Certain finite interval) ฟังก์ชั่นกำเนิดแฟคเตอร์ของโมเมนต์ของตัวแปรเชิงสุ่ม X กำหนดไว้ดังนี้

$$G(t) = E(t^X)$$

สำหรับ G(t) นี้ ถ้าหาอนุพันธ์เทียบกับ t และเมื่อให้ t = 1 เราจะได้โมเมนต์ผลคูณดังนี้

$$G^1(t) = E(Xt^{X-1})$$
$$G^1(1) = E(X)$$
$$G^2(t) = E[X(X-1)t^{X-2}]$$
$$G^2(1) = E[X(X-1)]$$

$$G^{k}(t) = E[X(X-1)\ldots(X-k+1)t^{x-k}]$$

$$G^{k}(1) = E[X(X-1)(X-2)\ldots(X-k+1)]$$

สำหรับ G(t) และ M(t) มีความสัมพันธ์กันดังนี้

$G(t) = M(\ln t)$; $M(t) = G(e^{t})$ โดยที่ $t^{x} = e^{x\ln t}$

G(t) นี้ยังได้ชื่อว่าฟังก์ชันกำเนิดความน่าจะเป็นของแต่ละค่าของตัวแปรเชิงสุ่มได้ ดังนี้ ให้ X เป็นตัวแปรเชิงสุ่มแบบไม่ต่อเนื่องและมีค่าเป็น 1, 2, 3, ..., n แล้ว

$$G(t) = E(t^{x})$$

$$= \sum_{x=1}^{n} t^{x}f(x)$$

$$= t^{1}f(1) + t^{2}f(2) + t^{3}(3) + \ldots + t^{n}f(n)$$

ดังนั้น $G^{1}(t) = f(1) + 2t\,f(2) + \ldots + nt^{n-1}f(n)$

$G^{2}(t) = 2f(2) + 3(2)tf(3) + \ldots + n(n-1)t^{n-2}f(n)$

$G^{n-1}(t) = (n-1)!\,f(n-1) + n!\,t\,f(n)$

$G^{n}(t) = n!\,f(n)$

เราจะเห็นว่า $G^{k}(0) = k!\,f(k)$; $k = 1, 2, 3, \ldots, n$. นั่นคืออนุพันธ์ ของ G(t) เมื่อคำนวณที่ t = o จะให้ผลคูณของค่าคงที่ค่าหนึ่งกับค่าของฟังก์ชันน่าจะเป็นสำหรับ X

ตัวอย่าง (1) **กำหนดให้** $f(x) = 1/4$; $x = 0, 2$

$= 2/4$; $x = 1$

เราจะได้ $G(t) = E(t^{x})$

$= 1/4 + 2/4t + 1/4t^{2} = 1/4(1+t)^{2}$

ดังนั้น $G(0) = 1/4 = f(0)$

$G^{1}(0) = 1/2 = f(1)$

$G^{2}(0) = 1/2 = 2!\,f(2)$; $f(2) = 1/4$

$G^{k}(t) = 0$; $k = 3, 4, 5, \ldots$

และ $G^{1}(1) = 1 = \mu_1' = \mu$

$G^{2}(1) = 1/2 = \mu_2' = E(x^{2}) - \mu$

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

$$G(t) = E(t^{x}) = \sum_{x=1}^{\infty} t^{x}\,1/2^{x}$$

$$= \sum_{x=1}^{\infty} (t/2)^{x} = t/(2-t) ; |t| < 2$$

เราจะได้ว่า $G^{1}(t) = 2/(2-t)^{2}$

$G^{2}(t) \quad = \quad 4/(2-t)^{3}$

$\vdots$

$G^{k}(t) \quad = \quad 2^{k}/(2-t)^{k+1} \quad k = 1, 2, 3, \ldots..$

จาก $G^{k}(o) \quad = \quad k!\,f(k)$ เราจะได้ว่า

$(1/k!)\,G^{k}(o) \quad = \quad f(k) = 1/2^{k};\ k = 1, 2, 3, \ldots..$

และ $G^{1}(1) \quad = \quad 2 = \mu$

$G^{2}(1) \quad = \quad 4 = E(X)^{2} - \mu$

(8) **ฟังก์ชันแสดงคุณลักษณะ** (Characteristic Function, C(t)) สำหรับ M(t) และ G(t) ที่กล่าวมานั้นของฟังก์ชันน่าจะเป็นของตัวแปรเชิงสุ่ม เราไม่สามารถหา C(t) และ G(t) ได้ แต่มีฟังก์ชันที่เรียกว่า ฟังก์ชันแสดงคุณลักษณะ C(t) นั้นสามารถหาได้ทุกฟังก์ชันน่าจะเป็นของตัวแปรเชิงสุ่ม และ C(t) นี้มีบทบาทสำคัญมากในทฤษฎีความน่าจะเป็นและการประยุกต์ทางสถิติ เช่น ในการแปลง (tranformation) ตัวแปรเชิงสุ่มต่าง ๆ หรือดูผลผลิตใหม่ (Reproduction) ของตัวแปรเชิงสุ่ม

นิยาม ตัวแปรเชิงสุ่ม X ใด ๆ ที่มีฟังก์ชันน่าจะเป็นนั้นจะเกี่ยวพันกับฟังก์ชันแสดงคุณลักษณะ C(t) ด้วยความสัมพันธ์ ดังนี้

เมื่อ t เป็นเลขจำนวนจริงและ i เป็นเลขจินตภาพ (imaginary unit) ที่กำหนดว่า $i = \sqrt{-1}$

สำหรับ e^{itx} นั้น เป็นฟังก์ชันค่าประกอบ (Complex - valued function) ที่กำหนดไว้ดังนี้

$$e^{itX} = \sum_{k=1} (itX)^{k}/k\,!$$

$$= \operatorname{Cos}(tX) + i \operatorname{Sin}(tX)$$

และ $\left| e^{itx} \right| = 1$ ดังนั้น C(t) จึงเป็นฟังก์ชันค่าประกอบด้วยและ C(t) จะผลิตโมเมนต์ต่าง ๆ เหมือนกับที่ผ่านมาสำหรับ C(t) จะมีคุณสมบัติที่น่าสนใจดังนี้

(ก) คำนวณ C(t) เมื่อ t = o จะได้ $C(0) = E(e^{1(o)X}) = E(e^{0}) = 1$

(ข) อนุพันธ์ของ C(t) ต่าง ๆ เมื่อเทียบ t จะเป็นดังนี้

$$C^{k(t)} = \frac{d^{k}}{dt^{k}}\left\{C(t)\right\} = (i)^{k}E(X^{k}e^{itX})$$

เมื่อคำนวณที่ t = o เราจะได้ $C^{k}(0) = (i)^{k}\,E(x)^{k}$

ซึ่งเราจะได้โมเมนต์ $E(X)^{k} = C^{k}(o)/(i)^{k}$

สำหรับ k = 1, 2, 3, 4 เราจะได้

$E(x) = C^{1}(0)/i,\ E(x)^{2} = C^{2}(0)/(-1)$

$E(X)^3 = C^3(0)/(-i)$, $E(x)^4 = C^4(0)$

(ค) ถ้าเราต้องการหาโมเมนต์รอบค่าเฉลี่ย μ เราจะพิจารณาจาก

$$C_{x-\mu}(t) = E[e^{it(X-\mu)}] = e^{-it\mu}C_x(t)$$

แล้วเราจะได้ $E(X-\mu)^k = C_{x-\mu}(t=0)/(i)^k$

คุณสมบัติของ C(t) เป็นหลักสำคัญในการศึกษาเรื่องการแจกแจงน่าจะเป็น ก็เพราะมันมีอยู่ฟังก์ชั่นเดียว (Uniques) สำหรับการแจกแจงน่าจะเป็นหรือตัวแปรเชิงสุ่มหนึ่งนั่นก็คือ

ถ้าสองฟังก์ชั่นน่าจะเป็น $f_x(t)$ และ $f_y(t)$ มี C(t) เหมือนกัน แล้วทั้งสองฟังก์ชั่นน่าจะเป็นจะเหมือนกัน

หรือ $C_x(t) = C_y(t) \iff f_x(t) = f_y(t)$

สำหรับทุกค่าจริงของ t

สำหรับตัวแปรเชิงสุ่ม X ถ้ากำหนด G(X) เป็นฟังก์ชั่นเชิงเส้น (linear) ของ X แล้วเราจะสามารถหา G(t) ของ (a + bX) ในเทอมของ C(t) ของ X ดังนี้

$$C_{a+bX}(t) = E(e^{it(a+bX)}) = e^{ita}E(e^{itbX})$$

$$= e^{ita}E(e^{itb})$$

ดังนั้น $C_{a+bX}(t) = e^{ita}C_x(tb)$

ตัวอย่าง **กำหนด** $f(x) = 0.5$; $n = -1, 1$

ดังนั้น $C(t) = \sum_x e^{itx}f(x) = 0.5e^{it(1)} + 0.5e^{it(+1)}$

$= 0.5(\cos t - i\sin t) + 0.5(\cos t + i\sin t) = \cos t$

$C(t=o) = \cos(0) = 1$

$E(X) = C^1(t=0)/i = -\sin(0)/i = 0$

$E(X^2) = C^2(t=0)/(i)^2 = \cos(0) = 1$

ทฤษฎีขีดจำกัดสู่ส่วนกลาง (Central Limit Theorem) ทฤษฎีนี้เป็นทฤษฎีสำคัญในทฤษฎีสถิติซึ่งกล่าวไว้ว่า

ถ้า $X_1, X_2, \ldots, X_n$ เป็นตัวแปรเชิงสุ่มที่เป็นอิสระกัน และต่างก็มีการแจกแจงเหมือนกันโดยมีค่าเฉลี่ยและความแปรปรวนเหมือนกันนั่นคือ $E(x_1) = E(x_2) = \ldots\ldots = E(x_n) = \mu$ และ $V(x_1 = V(x_2) = \ldots. = V(x_n) = \sigma^2$ แล้ว

$$\lim_{n\to\infty} P\left[a \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le b\right] = \int_a^b \frac{1}{\sqrt{2\pi}}e^{-z^2}dz$$

ในเมื่อ $X = \sum X_i/n$, $E(\bar{X}) = \mu$, $V(\bar{X}) = \sigma^2/n$,

และ $\frac{1}{\sqrt{2\pi}} e^{-z^2}$ เป็นฟังก์ชั่นหนาแน่นของการแจกแจงปกติมาตรฐาน N(O,I)

หรือกล่าวได้ว่า "การแจกแจงของ $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$ จะเป็นการแจกแจงปกติมาตรฐานถ้า n โตพอ"

ทฤษฎีนี้พิสูจน์ได้โดยฟังก์ชันแสดงคุณลักษณะ

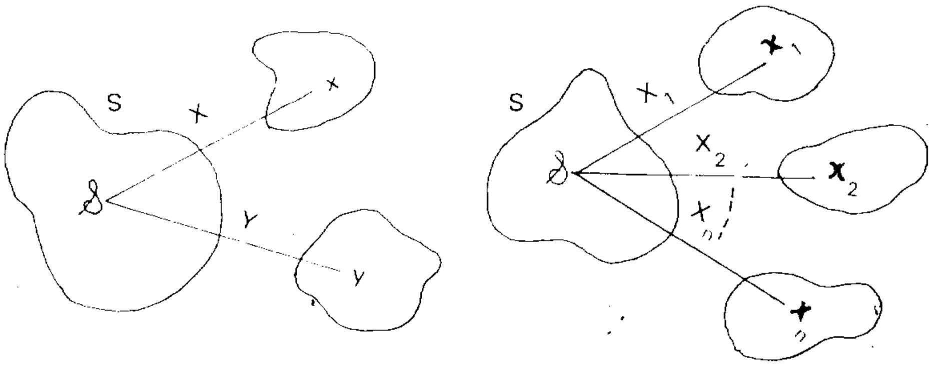
3.8 การแจกแจงหลายตัวแปร (Multivariate Distributions)

เท่าที่ผ่านมาเราพิจารณาผลทดลองของการลองเชิงสุ่มเพียงคุณลักษณะ (characteristic) เดียวเท่านั้น ตัวแปรเชิงสุ่มที่กำหนดในกลุ่มผลทดลอง (Sample space) เช่นนั้นเราเรียกว่า ตัวแปรเชิงสุ่มมิติเดียว (One-dimensional random variables) บ่อยครั้งที่เราสนใจผลทดลองมากกว่าหนึ่งคุณลักษณะ เช่น การวิเคราะห์คนงานที่ทำหน้าที่คุมเครื่องจักรนั้นจำเป็นจะต้องพิจารณาการศึกษา (E) และประสบการณ์ (X) ซึ่งเราจะได้ผลทดลอง (e,x) หรือผู้สมัครงานในบริษัท บริษัทอาจจะต้องศึกษาระดับสติปัญญา I และระดับการศึกษาภาคบังคับ E(formal education) ซึ่งจะได้ผลทดลองในตาราง (i,e) ในโอกาสที่เราทำการทดลองแล้วต้องวิเคราะห์ผลทดลองถึง 2,3,....,n คุณลักษณะ

ตัวแปรเชิงสุ่มที่กำหนดในกลุ่มผลทดลองเหล่านี้ เรียกว่าตัวแปรเชิงสุ่ม 2,3.....n มิติ ลองพิจารณานิยามต่อไปนี้

นิยาม ให้ S เป็นกลุ่มผลทดลองของการทดลองเชิงสุ่ม ถ้า X และ Y เป็นฟังก์ชั่นที่กำหนดจำนวนจริง X(S) และ Y(S) หรือ (x,y) ให้แก่แต่ละผลทดลอง $s \in S$ เราจะเรียก (X,Y) ว่าตัวแปรหรือเวคเตอร์เชิงสุ่มสองมิติ (Two-dimensional random Variable or vector)

ถ้า $X_1, X_2, \ldots, X_n$ เป็น n ฟังก์ชั่นที่กำหนดเลขจำนวนจริง $(x_1, x_2, \ldots, x_n)$ ให้แก่แต่ละผลทดลอง $s \in S$ เราเรียก $(X_1, X_2, \ldots, X_n)$ ว่าตัวแปร (เวคเตอร์) เชิงสุ่ม n มิติ



ในกรณีตัวแปรเดียว เราแบ่งตัวแปรออกเป็น 2 แบบ คือ ตัวแปรเชิงสุ่มต่อเนื่องและไม่ต่อเนื่อง ในกรณีหลายตัวแปรเราก็แบ่งได้เช่นกัน ดังนี้

นิยาม ตัวแปรเชิงสุ่มสองมิติ (X, Y) จะเป็นตัวแปรเชิงสุ่มไม่ต่อเนื่อง ถ้าค่าที่เป็นไปได้ของ (X, Y) มีจำกัดหรืออนันต์นับได้ หรือค่าที่เป็นไปได้ของ (X, Y) จะแทนได้เป็น (x_i, y_j); i, j = 1, 2, 3, ...

ในทำนองเดียวกัน $(X_1, X_2, ..., X_n)$ หรือ $\underline{X}$ เป็นตัวแปรเชิงสุ่มไม่ต่อเนื่อง n มิติถ้าค่าที่เป็นไปได้ของ $\underline{X}$ มีจำกัดหรืออนันต์นับได้

ตัวแปรเชิงสุ่ม t(X,Y) จะเป็นตัวแปรเชิงสุ่มต่อเนื่องสองมิติ ถ้า (X, Y) สามารถมีได้ทุกค่าในเซ็ทที่นับไม่ได้ของระนาบ (noncountable set of the Euclidean plane)

เช่นเดียวกัน $\underline{X}$ เป็นตัวแปรเชิงสุ่มต่อเนื่อง n มิติ ถ้าค่าที่เป็นไปได้ของ $\underline{X}$ สามารถมีได้ทุก ๆ ค่าในเขตของกลุ่ม (Space) n มิติ (Region of the n - dimensional space)

ในกรณีที่เราเกี่ยวข้องกับตัวแปรเชิงสุ่ม n มิติ เราอาจจะสนใจไม่เพียงแต่คุณลักษณะแต่ละอย่างแยกกันเท่านั้น ยังสนใจกับความสัมพันธ์ภายใน (interrelationships) ที่มีอยู่ระหว่างคุณลักษณะต่าง ๆ นั้น และการที่เราพิจารณาถึงความสัมพันธ์ระหว่างคุณลักษณะต่าง ๆ นั้นจำเป็นที่จะต้องศึกษาฟังก์ชั่นน่าจะเป็นของการเกิดร่วมกันของมันด้วย หรือการแจกแจงร่วม ซึ่งเราจะให้นิยามต่อไปนี้

นิยาม (ก) ให้ (X, Y) เป็นตัวแปรเชิงสุ่มไม่ต่อเนื่อง ฟังก์ชั่น f ที่กำหนดความน่าจะเป็น f(x, y) = P(X = x; Y = y) ให้แก่ผลทดลองที่เป็นไปได้ (x, y) นั้น เราจะเรียกว่า การแจกแจงน่าจะเป็นร่วมของ (X, Y) ถ้า f(x, y) สอดคล้องกับเงื่อนไขต่อไปนี้

(1) $f(x,y) \geqslant 0 \ ; \ \forall(x,y)$

(2) $\sum_{x,y} f(x,y) = 1$

ฟังก์ชั่น f จะเป็นการแจกแจงหลายตัวแปร (Multivariate distribution) ของตัวแปรเชิงสุ่มไม่ต่อเนื่อง ถ้า $f(\underline{x})$ กำหนดไว้ดังนี้

$$f(\underline{x}) = f(x_1, x_2, \ldots, x_n)$$
$$= P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$ และ $f(\underline{x})$ สอดคล้องกับเงื่อนไขต่อไปนี้

(1) $f(\underline{x}) \geqslant 0$ สำหรับทุก ๆ $(x_1, x_2, \ldots, x_n)$

(2) $\sum_{x_1}\sum_{x_2}\ldots\sum_{x_n} f(\underline{x}) = 1$

(ข) ให้ (X,Y) เป็นตัวแปรเชิงสุ่มต่อเนื่องที่มีค่าเป็นไปได้อยู่ในเขตของระนาบ R แล้ว ฟังก์ชันหนาแน่นน่าจะเป็นร่วม (Joint probability density function) f ของ (X,Y) เป็นฟังก์ชันที่สอดคล้องกับเงื่อนไขต่อไปนี้

(1) $f(x,y) \geqslant 0$ สำหรับทุก ๆ $(x,y) \in R$

(2) $\int_R\int f(x,y)\,dx\,dy = 1$

กรณีที่เป็นตัวแปรเชิงสุ่มหลายตัว $(X_1, X_2, \ldots, X_n)$ ที่มีค่าอยู่ในเขตของกลุ่ม (space) นั้น ฟังก์ชั่น f จะเรียกว่าการแจกแจงหลายตัวแปร ถ้าสอดคล้องกับเงื่อนไข ดังนี้

(1) $f(\underline{x}) \geqslant 0$ สำหรับทุก ๆ $\underline{x} \in \mathcal{V}$

(2) $\int\int_{\mathcal{V}}\ldots\int f(\underline{x})\,dx_1\,dx_2\ldots dx_n = 1$

สำหรับตัวแปรเชิงสุ่มแบบต่อเนื่องนี้เราจะกำหนด

$P\left\{x \leq X \leq x+\Delta x, \quad y \leq Y \leq y+\Delta y\right\}$

$= f(x,y)\,\Delta x \Delta y$ ถ้า Δx และ Δy เล็กพอ

หรือ $P[a \leq X \leq b, c \leq Y \leq d] = \int_a^b \int_c^d f(x,y)\,dx\,dy$ ถ้า $a < b$ และ $c < d$

ตัวอย่าง บริษัทต้องการเลือกตัวแทน 3 คน โดยการเลือกสุ่มจากผู้เข้าฝึกอบรมด้านบริหาร 10 คน เพื่อทำหน้าที่เป็นคณะกรรมการติดต่อของกลุ่มระหว่างระยะฝึกอบรม คนแรกที่สุ่มมาได้จะทำหน้าที่เป็นผู้ประสานงาน

ครึ่งหนึ่งของ 10 คนนี้เพิ่งจบจากมหาวิทยาลัย และยังไม่มีประสบการณ์มาก่อนเลย ที่เหลือเป็นพวกที่บริษัทจ้างไว้หลายปีแล้ว และกำลังจะได้รับแต่งตั้งเป็นนักบริหาร

การทดลองนี้จะได้กลุ่มผลทดลอง

$$S = \{NNN, NNO, NON, ONN, NOO, ONO, OON, OOO\}$$

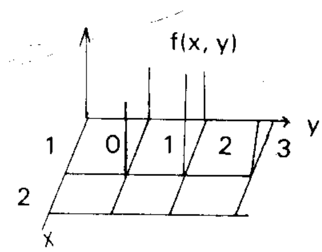
ในเมื่อ N เป็นพนักงานใหม่ และ O เป็นพนักงานเก่า

ถ้าตัวแปรเชิงสุ่ม X และ Y เป็นฟังก์ชั่นที่กำหนดในกลุ่มผลทดลองนี้โดยให้ X เป็นผู้ประสานงาน และให้ Y เป็นจำนวนพนักงานใหม่ในคณะกรรมการ แล้วเราสามารถแสดงผลทดลองค่าของ X กับ Y และความน่าจะเป็นที่สมนัยกับผลทดลองหรือค่าของ X และ Y ได้ดังนี้

ผลทดลอง (s_i)	X	Y	ความน่าจะเป็น P(s_i)
NNN	1	3	$(\frac{5}{10})(\frac{4}{9})(\frac{3}{8}) = \frac{3}{36}$
NNO	1	2	$(\frac{5}{10})(\frac{4}{9})(\frac{5}{8}) = \frac{5}{36}$
⋮		⋮	
OOO	0	0	$(\frac{5}{10})(\frac{4}{9})(\frac{3}{8}) = \frac{3}{36}$

จากตารางข้างบนเราสามารถสร้างฟังก์ชันน่าจะเป็นร่วมของ (X, Y) โดยตาราง ได้ ดังนี้

X \ Y	0	1	2	3	
0	$\frac{3}{36}$	$\frac{10}{36}$	$\frac{5}{36}$	0	$\frac{18}{36}$
	0	$\frac{5}{36}$	$\frac{10}{36}$	$\frac{3}{36}$	$\frac{18}{36}$
	$\frac{3}{36}$	$\frac{15}{36}$	$\frac{15}{36}$	$\frac{3}{36}$	1



ตัวอย่าง ต้นทุนและกำไรเป็นตัวแปรเชิงสุ่มร่วมที่มีฟังก์ชันเป็น

$f(x,y) = 4xy; \quad 0 \leq x, y \leq 1$

ฟังก์ชันนี้เป็นฟังก์ชันแจกแจงร่วมเพราะ $\int_0^1\int_0^1 4xy\,dx\,dy = \int_0^1 4y\,(x^2/2)\Big|_0^1\,dy$

$$= \int_0^1 (4/2)\,y\,dy = 4y^2/2(2)\Big]_0^1 = 1$$

สำหรับความน่าจะเป็นเมื่อ $x \geq \frac{1}{2}$ และ $y \geq 0$ หาได้ดังนี้

$P[X \geq \frac{1}{2} Y \geq 0] = \int_{\frac{1}{2}}^1\int_0^1 4xy\,dx\,dy = 3/4$

3.8.1 **ฟังก์ชันแจกแจงสะสมของตัวแปรเชิงสุ่มหลายตัว** (Cumulative Distribution Functions, CDF.) ในตัวแปรเชิงสุ่มมิติเดียวฟังก์ชันแจกแจงสะสมมีบทบาทสำคัญอยู่มาก สำหรับตัวแปรเชิงสุ่มหลายตัว เราจะนิยมฟังก์ชันแจกแจงสะสม ดังนี้

นิยาม : ให้ (X,Y) เป็นตัวแปรเชิงสุ่มสองมิติ และ (x,y) เป็นเลขจำนวนจริงแล้วฟังก์ชัน F ที่กำหนดความน่าจะเป็น F(x, y) ให้แก่เหตุการณ์ที่ **X** มีค่าน้อยกว่าหรือเท่ากับ x และ **Y** มีค่าน้อยกว่าหรือเท่ากับ y นั้น จะเรียกว่า **ฟังก์ชันแจกแจงสะสมของตัวแปรเชิงสุ่ม** (X, Y)

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

$$= \begin{cases} \sum_{X \leq x} \sum_{Y \leq y} f(x, y); & x, y \text{ ไม่ต่อเนื่อง} \\ \int_{X \leq x} \int_{Y \leq y} f(x, y)\,dx\,dy; & x, y \text{ ต่อเนื่อง} \end{cases}$$

สำหรับตัวแปรเชิงสุ่มหลายตัว $(X_1, X_2, \ldots, X_n)$ ฟังก์ชันแจกแจงสะสม F กำหนดไว้ดังนี้

$F(\underline{X}) = (x_1, x_2, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)$

ตัวอย่าง จากฟังก์ชันร่วมน่าจะเป็น

$f(x, y) = (3 - x - y)/3;\ 0 \leq x \leq 2,\ 0 \leq y \leq 1$

เราหา $F(x, y)$ ได้ดังนี้

$$F(x, y) = \int_0^x \int_0^y (1/3)(3 - x - y)\,dx\,dy$$

$$= 1/3\left(3xy - \frac{xy^2}{2} - \frac{x^2y}{2}\right)$$

ดังนั้น $P(X \leq 2, Y \leq 1) = F(2, 1) = 1$

และ $P(X \leq 1, Y \leq 1) = F(1, 1) = \frac{2}{3}$

ถ้า F เป็นฟังก์ชันแจกแจงสะสมของตัวแปรเชิงสุ่มสองมิติที่มีฟังก์ชันหนาแน่นน่าจะเป็นร่วม f แล้ว

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

ในเมื่อ F หาอนุพันธ์ได้

โดยปกติตัวแปรเชิงสุ่มแบบต่อเนื่องสองมิติเรามักจะใช้ฟังก์ชันการแจกแจงอธิบายตัวแปรเชิงสุ่มนั้น

3.8.2 ฟังก์ชันน่าจะเป็นทางเดียวและแบบเงื่อนไข (Marginal and conditional Probability Functions) ถึงแม้ว่าเราจะมีฟังก์ชันน่าจะเป็นร่วมสำหรับ (x, y) เราก็ยังสนใจที่จะหาฟังก์ชันของตัวแปรเชิงสุ่มแต่ละตัวแยกกัน และฟังก์ชันของแต่ละตัวจากตัวแปรเชิงสุ่มที่มีการแจกแจงร่วม เราจะเรียกว่าฟังก์ชันน่าจะเป็นทางเดียว (Marginal Probability function) ซึ่งจะนิยามดังต่อไปนี้

นิยาม เมื่อ f เป็นฟังก์ชันน่าจะเป็นร่วมของตัวแปรเชิงสุ่ม (X, Y) แล้ว g และ h จะเป็นฟังก์ชันน่าจะเป็นทางเดียวของ X และ Y ถ้านิยามไว้ดังนี้

$$g(x) = \int_{\forall y} f(x, y)\,dy$$

$$= \sum_{\forall y} f(x, y)$$

และ $$h(y) = \int_{\forall x} f(x, y)\,dx$$

$$= \sum_{\forall x} f(x, y)$$

สำหรับตัวแปรเชิงสุ่ม n มิติ $(X_1, X_2, \ldots, X_n)$ ที่มีการแจกแจง f เราจะหาฟังก์ชันน่าจะเป็นทางเดียว g ของ $X_1, X_2, \ldots,$ และ X_n ได้ดังนี้

$$g(x_i) = \begin{cases} \sum\sum \cdots \sum f(x_1, x_2, \ldots, x_n) \\ \text{ยกเว้น } x_i \\ \iint \cdots \int f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n \\ \text{ยกเว้น } x_i \end{cases}$$

และเราสามารถหาฟังก์ชันน่าจะเป็นทางเดียวสำหรับกลุ่มของตัวแปร เช่น (X_1, X_2) หรือ (X_3, X_5, X_7) ก็ได้ ดังนี้

$$g(x_1, x_2) = \int\int_{\text{ยกเว้น } x_1, x_2} \cdots \int f(\underline{x})\, dx_3\, dx_4 \ldots dx_n$$

ตัวอย่าง (1) ถ้า X และ Y เป็นตัวแปรเชิงสุ่มร่วมที่มีฟังก์ชัน ดังนี้

$$f(x, y) = \frac{1}{4}; (x, y) = (-1, -1), (-1, 1), (1, -1), (1, 1)$$

แล้ว $g(x) = \frac{1}{2}; X = -1, 1$ และ $h(y) = \frac{1}{2}; y = -1, 1$

(2) ฟังก์ชันน่าจะเป็นร่วมของ (X, Y) กำหนดว่าเป็น

$$f(x, y) = (6 - x - y)/8; 0 \leq x \leq 2, 2 \leq y \leq 4$$

แล้ว $g(x) = \int_2^4 \frac{1}{8}(6 - x - y)\, dy = \frac{1}{4}(3 - x)$

และ $h(y) = \int_0^2 \frac{1}{8}(6 - x - y)\, dx = \frac{1}{4}(5 - y)$

สำหรับตัวแปรเชิงสุ่มร่วม (X, Y), เมื่อเราพิจารณาฟังก์ชันน่าจะเป็นทางเดียวของ X นั้น เราก็พิจารณาเฉพาะค่าฟังก์ชันของ X โดยไม่สนใจหรือเกี่ยวข้องกับค่าของ Y แต่ถ้าเราสนใจกับค่าอื่น ๆ ที่เจาะจงไว้หรือ Y = y ฟังก์ชันน่าจะเป็นของ X เมื่อกำหนด Y = y เราจะเรียกว่า ฟังก์ชันน่าจะเป็นเงื่อนไขของ X สำหรับ Y = y ที่กำหนดไว้ (Conditional P.D.F. of X for given Y = y) ดังนิยามต่อไปนี้

นิยาม ให้ (X, Y) เป็นตัวแปรเชิงสุ่มสองมิติที่มีฟังก์ชันน่าจะเป็นร่วม f และให้ g และ h เป็นฟังก์ชันน่าจะเป็นทางเดียวของ X และ Y ตามลำดับแล้วฟังก์ชันน่าจะเป็นเงื่อนไขของ X กำหนด Y = y, f(x/y) นิยามไว้ดังนี้

$$f(x/y) = f(x, y)/h(y), h(y) > 0$$

และฟังก์ชันน่าจะเป็นเงื่อนไขของ Y กำหนด X = x, g(y/x) นิยามไว้ดังนี้

$$f(y/x) = f(x, y)/g(x) ; g(x) > 0$$

จากนิยามข้างบนเราจะได้อีกว่า

$$f(x, y) = \begin{cases} g(x/y)\, h(y) \\ g(y/x)\, g(x) \end{cases}$$

ตัวอย่าง (1) ตารางต่อไปนี้จะแสดงฟังก์ชันน่าจะเป็นร่วมของ (X, Y), **ฟังก์ชันน่าจะเป็นทาง**เดียวของ x และ y และฟังก์ชันน่าจะเป็นเงื่อนไข

	Y	y_1	y_2	y_3	$f(x_i)$
X	x_1	.05	.10	.20	.35
	x_2	.20	.20	.05	.45
	x_3	.05	.05	.10	.20
	$f(y_j)$	.30	.35	.35	1.00

	$f(x_i/y_1)$	$f(x_i/y_2)$	$f(x_i/y_3)$
x_1	1/6	2/7	4/7
x_2	4/6	4/7	1/7
x_3	1/6	1/7	2/7

	y_1	y_2	y_3
$f(y_j/x_1)$	1/7	2/7	4/7
$f(y_j/x_2)$	4/9	4/9	1/9
$f(y_j/x_3)$	1/4	1/4	2/4

$f(y_j/x_i)$

$f(x_i/y_j)$

$$f(x_1/y_1) = g(x_1,y_1)/h(y_1) = .05/.30 = 1/6$$
$$f(x_2/y_1) = g(x_2,y_1)/h(y_1) = .20/.30 = 4/6$$
$$f(y_1/x_1) = .05/.35 = 1/7$$
$$f(y_1/x_2) = .10/.35 = 2/7$$
$$\vdots$$
$$f(y_3/x_3) = .10/.20 = 1/2$$

$$f(x_1/y_3) = \frac{.20}{.35} = 4/7$$
$$f(x_2/y_3) = \frac{.05}{.35} = 1/7$$
$$f(x_3/y_3) = \frac{.10}{.35} = 2/7$$

ตัวอย่าง (1) จาก $f(x, y) = \frac{1}{8}(6-x-y)\ ;\ 0 \leq x \leq 2,\quad 2 \leq y \leq 4$

เราจะได้ $g(x) = \frac{1}{4}(3-x)$

$h(y) = \frac{1}{4}(5-y)$

ดังนั้น $f(x/y) = f(x, y)/h(y) = \frac{1}{8}(6-x-y)/\frac{1}{4}(5-y)$

$= (6-x-y)/2(5-y)$

และ $f(y/x) = f(x, y)/g(x) = \frac{1}{8}(6-x-y)/\frac{1}{4}(3-x)$

$= (6-x-y)/2(3-x)$

สำหรับตัวแปรเชิงสุ่มหลายตัวเราก็หาฟังก์ชั่นน่าจะเป็นเงื่อนไขได้ทำนองเดียวกัน เช่น

$f(x_1/x_2, x_3, \ldots, x_n) = f(\underline{x})/f(x_2, x_3, \ldots x_n)$

หรือ $f(x_1, x_2, \quad x_i/x_{i+1}, x_{i+2}, \ldots, x_n) = f(\underline{x})/f(x_{i+1}, \ldots, x_n)$

3.8.3 การเป็นอิสระทางสถิติของตัวแปรเชิงสุ่ม (Statistical Independence of random variables) ในเรื่องของความน่าจะเป็นเบื้องต้น เราให้นิยามการเป็นอิสระระหว่างสองเหตุการณ์ A และ B ไว้แล้ว แนวความคิดในเรื่องการเป็นอิสระของตัวแปรเชิงสุ่มของตัว X และ Y ก็เช่นเดียวกัน คือ เราจะพูดว่า X เป็นอิสระจาก Y ถ้าผลทดลองของ X ไม่มีอิทธิพลต่อผลทดลองของ Y **ตัวอย่างเช่น** ถ้า X มีค่าเป็น 1, 2, ..., 6 สำหรับผลทดลองที่เป็นไปได้เมื่อทอดลูกเต๋า และ Y มีค่าเป็น 0 และ 1 สำหรับผลทดลองที่เป็นก้อยและหัวในการโยนเหรียญแล้ว (X, y) จะแจกแจงร่วม แต่จะเป็นอิสระกัน (ในแง่ของความน่าจะเป็น ; นั่นคือ $f(4, 0) = P(X = 4, \quad Y = 0) = P(X = 4)\,P(Y = 0)$ $= f(4).f(0)$ เมื่อ (x, y) ที่เป็นไปได้จะทำให้ $f(x, y) = f(x)\,f(y)$ เป็นจริงเสมอแล้ว X และ Y จะเรียกว่าเป็นอิสระกันในเชิงสถิติ (statistically independent) แต่ถ้าไม่เป็นจริงเสมอแล้ว เราจะเรียก X และ Y ว่าพึ่งพิงกันในเชิงสถิติ (Statistically Dependent) เราจะให้นิยามการเป็นอิสระดังนี้

นิยาม ให้ (X, Y) เป็นตัวแปรเชิงสุ่มของมิติแล้ว (X, Y) จะเรียกว่าตัวแปรเชิงสุ่มที่เป็นอิสระกันก็ต่อเมื่อ $f(x, y) = g(x)\,h(y)$ สำหรับทุก ๆ (x, y)

สำหรับตัวแปรเชิงสุ่มหลายตัว $(X_1, X_2, \ldots, X_n)$ จะเป็นอิสระกันก็ต่อเมื่อ $f(X_1, X_2, \ldots, X_n) = f(x_1)\,f(x_2)\ldots\ldots f(x_n)$

หรือ $f(\underline{x}) \;=\; \prod_{i=1}^{n} f(x_i)$

จากนิยามของฟังก์ชั่นน่าจะเป็นเงื่อนไข เราจะได้ทฤษฎีเกี่ยวกับการเป็นอิสระของตัวแปรได้ดังนี้

ทฤษฎี ให้ (X, Y) เป็นตัวแปรเชิงสุ่มสองมิติแล้ว X และ Y จะเป็นอิสระกันก็ต่อเมื่อ

$f(x/y) \;=\; f(x)$

หรือ $f(y/x) \;=\; f(y)$

ตัวอย่าง (1) เครื่องจักรตัวหนึ่งสำหรับใช้ในงานเฉพาะอย่างหนึ่งในตอนเช้า และสำหรับงานอย่างอื่นในตอนบ่าย ให้ X และ Y แทนจำนวนครั้งที่เครื่องหยุดเสียในตอนเช้าและบ่ายตามลำดับ โดยที่ (X, Y) มีการแจกแจงร่วม ดังนี้

y \ x	0	1	2	h(y)
0	0.1	0.2	0.2	0.5
1	0.04	0.08	0.08	0.2
2	0.06	0.12	0.12	0.3
g(x)	0.2	0.4	0.4	1.0

จากการคำนวณสำหรับทุก ๆ (x, y) เราจะได้

$$f(x, y) = g(x)\,h(y)$$

เช่น $f(0, 1) = f(0)\,f(1)$ นั่นคือ $0.04 = (0.2)(0.2)$

ดังนั้น X และ Y เป็นตัวแปรเชิงสุ่มที่เป็นอิสระกัน

(2) ให้ X และ Y เป็นอายุของเครื่องมืออีเล็คโทรนิคส์ 2 อย่าง สมมติว่ามี การแจกแจงร่วมเป็นดังนี้

$$f(x, y) = e^{-(x+y)};\ x > 0,\ y > 0$$

เราจะพบว่า $f(x) = e^{-x}$ และ $f(y) = e^{-y}$

และ $f(x/y) = f(x, y)/f(y) = e^{-(x+y)}/e^{-y}$

$= e^{-x}$

หรือ $f(y/x) = e^{-y}$

ดังนั้น X และ Y เป็นตัวแปรเชิงสุ่มที่เป็นอิสระกันในเชิงสถิติ

3.8.4 ค่าคาดหวังของฟังก์ชั่นของสองตัวแปรเชิงสุ่ม (Expected Values of a function of two random variables) เราสามารถใช้ฟังก์ชั่นน่าจะเป็นร่วมของตัวแปรเชิงสุ่ม หาค่าคาดหวังของฟังก์ชั่นที่มีตัวแปรเชิงสุ่มสองตัวหรือมากกว่าได้

สมมติว่า X และ Y เป็นตัวแปรเชิงสุ่มและ h (X, Y) กำหนดฟังก์ชั่นของ X และ Y เช่น h(X, Y) อาจจะเป็น $Z = X + Y$, $Z = XY$ หรือ $Z = (X - Y)^2$ เป็นต้น ค่าคาดหวังของ h(X, Y) กำหนดไว้ดังนี้

นิยาม ให้ (X, Y) เป็นตัวแปรเชิงสุ่มสองมิติ และ h(X, Y) เป็นฟังก์ชั่นของ X และ Y แล้วค่าคาดหวังของ h(X, Y) กำหนดไว้ดังนี้

$$E[h(X, Y)] = \begin{cases} \sum_x \sum_y h(x, y)\, f(x, y) \\ \int_x \int_y h(x,y)\, f(x,y)\, dx\, dy \end{cases}$$

ทั้งนี้ E [h(X, Y)] ต้องหาค่าได้ (absolutely convergent)

ตัวอย่าง สามีภรรยาคู่หนึ่งเป็นนักเศรษฐศาสตร์และบริหารธุรกิจ ให้ X, Y เป็นรายได้ของสามีและภรรยาเมื่อปีที่แล้ว จากประสบการณ์เขาทั้งสองพบว่า ฟังก์ชั่นการแจกแจงร่วมของรายได้ (10,000 บาท) เป็นดังนี้

$$f(x, y) = 1/15;\quad 10 < x < 15,\ 8 < y < 11$$

ดังนั้นค่าคาดหวังของรายได้รวมจะเป็น

$$E(X + Y) = \int_8^{11} \int_{10}^{15} (x + y)(1/15)\, dx\, dy = 22$$

และค่าคาดหวังของผลต่างระหว่างรายได้ของสามีกับภรรยาจะเป็น

$$E(X - Y) = \int_8^{11} \int_{10}^{15} (x - y)(1/15)\, dx\, dy = 3$$

ตัวอย่าง สถานีบริการสามทหารขายน้ำมันเบนซินได้เฉลี่ย 500 แกลลอนต่อวัน สถานีบริการจะได้กำไรแกลลอนละ 25 สตางค์ เจ้าของสถานีบริการต้องการประมาณกำไรทั้งหมดจากการขายใน 100 วันข้างหน้า โดยสมมติว่าการขายแต่ละวันเป็นอิสระกัน เขาจะประมาณได้เท่าไร

ให้ X_i เป็นตัวแปรเชิงสุ่มของน้ำมันที่ขายได้แต่ละวัน แล้ว $E(X_i) = 500$

ให้ Y_i เป็นตัวแปรเชิงสุ่มของกำไรแต่ละแกลลอนที่ขายได้ แล้ว $Y_i = 25\,X_i$ นั่นคือกำไรคาดหวังของแต่ละวันจะเป็น $E(Y_i) = 25(500)$

และให้ W เป็นตัวแปรเชิงสุ่มของกำไรรวมสำหรับ 100 วันข้างหน้า แล้ว $W = Y_1 + Y_2 + \ldots + Y_{100}$

ดังนั้น ตามทฤษฎี $E(W) = E(Y_1) + E(Y_2) + \ldots + E(Y_{100})$

$= 100\,(25)\,500 = 12500$ บาท

สำหรับฟังก์ชั่นน่าจะเป็นเงื่อนไขนั้น คุณลักษณะที่สำคัญของมันก็คือ ค่าคาดหวังเงื่อนไข (Conditional Expectation) ซึ่งกำหนดไว้ดังนี้

นิยาม ค่าคาดหวังเงื่อนไขของตัวแปรเชิงสุ่ม X เมื่อกำหนด Y = y กำหนดไว้ดังนี้

$$E(X/Y = y) = \begin{cases} \sum\limits_{\forall x} x\, f(x/y) \\ \int\limits_{\forall x} x\, f(x/y)\, dx \end{cases}$$

ตัวอย่าง กำหนด $f(x, y) = 8xy\,;\ 0 < y < x < 1$

เราจะได้ $g(x) = \int_0^x 8xy\, dy = 4x^3;\ 0 < x < 1$

และ $f(y/x) = f(x, y)/g(x) = 8xy/4x^3$

$= 2y/x^2\,;\ 0 < x < 1,\ 0 < y < x$.

ดังนั้น $E(Y/x) = \int_0^x y\, f(y/x)\, dy = \int_0^x y\left(\frac{2y}{x^2}\right) dy = 2/3\,x\,;\ 0 < x < 1$

จากนิยามและตัวอย่างเราจะเห็นได้ว่า $E(Y/x)$ เป็นฟังก์ชั่นของ X ถ้ากำหนดให้เป็น h(x) นั่นคือ

$E(Y/x) = g(x)$ ฉะนั้น เราสามารถจะหาค่าคาดหวังของตัวแปรเชิงสุ่ม g(x) ได้ดังทฤษฎีต่อไปนี้

ทฤษฎี ถ้า (X, Y) เป็นตัวแปรเชิงสุ่มสองมิติ แล้วเราจะได้ว่า

$E[E(Y/X)] = E(X)$

และ $E[E(X/Y)] = E(Y)$

ตัวอย่าง จำนวนของที่ผู้ผลิตส่งให้ร้านขายแต่ละวันเป็นตัวแปรเชิงสุ่ม N ซึ่งมีการแจกแจงดังนี้

n	10	11	12	13	14	15
f(n)	.05	.10	.10	.20	.35	.20

โอกาสที่ของชิ้นหนึ่งชิ้นใดจะเสียเท่ากันหมด คือ 0.10

ถ้า X เป็นจำนวนของที่เสียที่เจ้าของร้านไม่รับในวันหนึ่ง จงพิจารณา E(X)

เมื่อพิจารณา N = n เราจะได้ว่า X จะมีการแจกแจงแบบทวินาม นั่นคือ $f(x/n) = \binom{n}{x} (0.10)^x (0.90)^{n-x}$; x = 0, 1, 2, ..., n โดยที่ N ก็เป็นตัวแปรเชิงสุ่ม เราจึงหา E(X) ได้ดังนี้

$$E(X) = E[E(X/N)]$$

แต่ $E(X/N = n) = \Sigma x f(x/n)$

$$= \Sigma_x x \binom{n}{x} (0.10)^x (0.90)^{n-x} = 0.10n$$

ดังนั้น $E(X/N) = 0.10N$

นั่นคือ $E(X) = E[E(X/N)] = E(0.10N) = 0.10E(N)$

แต่ $E(N) = 10(0.05) + 11(0.10) + \ldots + 15(0.20) = 13.3$

เพราะฉะนั้น $E(X) = 0.10(13.3) = 1.33$

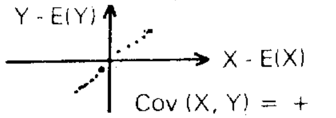
3.8.5 ความแปรปรวนร่วม (Covariance) สำหรับสองตัวแปรเชิงสุ่ม X และ Y นั้น เราไม่เพียงแต่ต้องการทราบว่ามันขึ้นอยู่กับกันหรือพึ่งพิงกัน (Statistically dependent) หรือไม่เท่านั้น แต่เรายังต้องการทราบว่ามันพึ่งพิงขนาดไหน (strength of their dependence) สำหรับมาตราวัดที่เราใช้ชี้ขนาดหรือระดับของการพึ่งพิงเชิงเส้น (linear dependence) เรียกว่า ความแปรปรวน (Covariance) ระหว่างสองตัวแปรเชิงสุ่ม

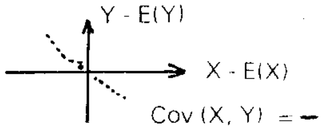
ดังนั้นจึงกล่าวได้ว่า ความแปรปรวนร่วมเป็นมาตราวัดดีกรีหรือความมากน้อยที่ตัวแปรสองตัวมีสัมพันธ์เชิงเส้น (linealy related) กัน นั่นก็คือ เป็นตัวชี้ (indicator) ถึงดีกรีที่สองตัวแปรเคลื่อนไปด้วยกัน (move together) ในแบบเชิงเส้น

มาตราวัดนี้จะเป็นอิสระกับจุดกำเนิด (origin) จึงสามารถย้ายจุดกำเนิด (0, 0) ไปไว้ที่ (E(X), E(Y)) ได้ แล้วเราจะได้ตัวแปรใหม่เป็น X - E(X) และ Y - E(Y) และค่าคาดหวังของผลคูณระหว่างตัวแปรใหม่ เรียกว่าความแปรปรวนร่วมของ X และ Y ดังนิยามต่อไปนี้

นิยาม (X, Y) เป็นตัวแปรเชิงสุ่มสองมิติ ความแปรปรวนร่วมของ X และ Y, Cov (X, Y) กำหนดไว้ดังนี้

$$\text{Cov}(X, Y) = E[X - E(X)][Y - E(Y)]$$
$$= E(XY) - E(X)E(Y)$$





ถ้า X มีแนวโน้มไปทางมาก เมื่อ Y มาก และไปทางน้อยเมื่อ Y น้อย แล้วจุด (x, y) จะอยู่ในส่วนที่ 1 หรือ 3 ซึ่ง Cov (X, Y) จะเป็นบวก (ตามรูปข้างบน) แต่ถ้า X มีแนวโน้มไปทางมาก เมื่อ Y น้อย และไปทางน้อยเมื่อ Y มาก จุด (x, y) จะอยู่ในส่วนที่ 2 หรือ 4 แล้ว Cov (X, Y) จะเป็นลบ

ถ้าไม่มีแนวโน้มที่ X และ Y จะเคลื่อนไปด้วยกัน แล้ว Cov (X, Y) จะใกล้ ๆ ศูนย์ จากความจริงที่ว่า ตัวแปรเชิงสุ่ม X และ Y ที่เป็นอิสระกันนั้น E(X, Y) = E(X) E(Y) ดังนั้น Cov (X, Y) = 0 ซึ่งเราจะได้เป็นทฤษฎีดังนี้

ทฤษฎี ให้ X และ Y เป็นตัวแปรเชิงสุ่มของตัวที่มีการแจกแจงร่วมกัน ถ้า X และ Y เป็นอิสระกันในเชิงสถิติแล้ว Cov (X, Y) = 0

ทฤษฎีเพียงแต่กล่าวว่า ถ้า X และ Y เป็นอิสระกันแล้ว ความแปรปรวนร่วมระหว่างสองตัวแปรเชิงสุ่มจำเป็นต้องเท่ากับศูนย์ แต่ทฤษฎีไม่หมายความถึงว่า ถ้าความแปรปรวนร่วมระหว่างสองตัวแปรเชิงสุ่มใด ๆ เป็นศูนย์แล้ว สองตัวแปรนั้นจะเป็นอิสระในเชิงสถิติ ซึ่งสองตัวแปรเชิงสุ่มนั้นอาจจะไม่พึ่งพิงกันเชิงเส้นก็ได้ ถึงแม้ว่าความแปรปรวนระหว่างสองตัวแปรจะเป็นศูนย์ ลองพิจารณาตัวอย่างต่อไปนี้

ตัวอย่าง ฟังก์ชันน่าจะเป็นร่วมของสองตัวแปรเชิงสุ่ม (X, Y) กำหนดไว้ดังนี้

x \ y	1	2	3	f(x)
1	0	$\frac{1}{3}$	0	$\frac{1}{3}$
2	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
f(y)	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1.00

จะได้ว่า $E(X) = 1(\frac{1}{3}) + 2(\frac{2}{3}) = \frac{5}{3}$

$$E(Y) = 1(\frac{1}{3}) + 2(\frac{1}{3}) + 3(\frac{1}{3}) = \frac{6}{3} = 2$$

$$E(XY) = 1(1)0 + 1(2)\frac{1}{3} + (13)\,0 + 2(1)\frac{1}{3} + 2(2)\,0$$

$$+ 2(3)\frac{1}{3} = 10/3$$

ดังนั้น $Cov(x, Y) = E(XY) - E(X)E(Y)$

$$= 10/3 - 5/3\ (2) = 0$$

แต่ X และ Y ไม่เป็นอิสระกัน เพราะ $f(x, y) \neq f(x) f(y)$

กฎเกี่ยวกับความแปรปรวนและความแปรปรวนร่วม ถ้าให้ X และ Y เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงร่วมกัน และ a, b, c, d เป็นค่าคงที่แล้ว

(1) $Cov(X + a, Y + b) = Cov(X, Y)$

(2) $Cov(aX + b, cY + d) = ac\, Cov(X, Y)$

(3) $V(aX \pm bY) = a^2V(X) + b^2V(Y) \pm 2ab\, Cov(X, Y)$

ถ้า X และ Y เป็นอิสระกันแล้ว $V(aX \pm bY) = a^2V(X) + b^2V(Y)$

3.8.6 **สัมประสิทธิ์สหสัมพันธ์** (Correlation coefficient) ค่าของความแปรปรวนร่วมจะมีหน่วยเป็นผลคูณของหน่วยของตัวแปรทั้งสอง บางครั้งหน่วยทั้งสองก็แตกต่างกัน เช่น X เป็นตัวแปรเชิงสุ่มเกี่ยวกับน้ำหนักมีหน่วยเป็นปอนด์ และ Y เป็นส่วนสูงมีหน่วยเป็นฟุต นักสถิติจึงพยายามหลีกเลี่ยงด้วยการกำจัดหน่วย โดยการทำตัวแปรให้อยู่ในรูปตัวแปรมาตรฐาน หรือโดยเอาส่วนเบี่ยงเบนของตัวแปรทั้งสองซึ่งมีหน่วยเช่นเดียวกับตัวแปรเชิงสุ่ม ไปหารความแปรปรวนร่วมก็จะได้ตัวเลขจริง (Pure number) ดังนั้น เราจึงได้มาตราวัดความสัมพันธ์เชิงเส้นของตัวแปรเชิงสุ่มสองตัวด้วย สัมประสิทธิ์สหสัมพันธ์ ρ

ซึ่งนิยามไว้ดังนี้

นิยาม ให้ X และ Y เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงร่วมสองตัวสัมประสิทธิ์ สหสัมพันธ์ระหว่าง X และ Y กำหนดไว้ว่า

$$\rho_{xy} = \frac{Cov(X, Y)}{\sigma_x \sigma_y} = E\left(\frac{X - E(X)}{\sigma_x}\right)\left(\frac{Y - E(Y)}{\sigma_y}\right)$$

สำหรับตัวแปรเชิงสุ่ม X และ Y ที่เป็นอิสระกันนั้น $Cov(X, Y) = 0$ ดังนั้น จึงไม่มีความสัมพันธ์ระหว่างสองตัวแปร นั่นคือ $\rho = 0$

เมื่อตัวแปรสองตัวมีความผันแปรร่วมอย่างสมบูรณ์ (Perfect covariability) และผันแปรไปทางเดียวแล้ว $\rho = 1$ แต่ถ้าผันแปรไปทิศทางตรงกันข้ามแล้ว $\rho = -1$

ดีกรีของความผันแปรที่น้อยกว่าจะให้สัมประสิทธิ์สหสัมพันธ์มีค่าเป็น $0 < \rho < 1$ หรือ $-1 < \rho < 0$ นั่นคือ $-1 < \rho < 1$

ทฤษฎี (1) สำหรับตัวแปรเชิงสุ่ม x และ y ใด ๆ เราจะได้ว่า $-1 < \rho < +1$

(2) ถ้า x และ y ต่างก็เป็นตัวแปรเชิงสุ่มที่มีความสัมพันธ์ในรูปฟอร์ม $Y = Ax + B$ โดยที่ A และ B เป็นค่าคงที่ แล้ว $\rho^2 = 1$ นั่นคือ $\rho = +1$ ถ้า $A > 0$ หรือ $\rho = -1$ ถ้า $A < 0$

(3) ถ้า ρ_{xy} เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรเชิงสุ่ม X และ Y และถ้า V = AX + B และ W = Cy + D โดยที่ A, B, C, D เป็นค่าคงที่ แล้ว

$$\rho_{vw} = \frac{AC}{|AC|}\rho_{xy}$$

นั่นคือ $\rho_{vw} = \rho_{xy}$ ถ้า A และ C มีเครื่องหมายเหมือนกัน หรือ $\rho_{vw} = -\rho_{xy}$ ถ้า A และ C มีเครื่องหมายต่างกัน

ตัวอย่าง ให้ X และ Y เป็นตัวแปรเชิงสุ่มเกี่ยวกับเปลี่ยนแปลงราคาประจำวันของ 2 หุ้น (ก, ข) และมีฟังก์ชั่นน่าจะเป็นร่วมดังนี้

x \ y	-2	2
-2	.4	1
2	.1	4

สำหรับ (X, Y) เราจะพบว่าทุก ๆ (x, y) **นั้น** $f(x, y) \neq f(x)\, f(y)$ **เช่น** $f(-2, 2) = 0.4$ และ $f(x = 2)\, f(y = 2) = 0.5(0.5) = 0.25$

ดังนั้น X และ Y พึ่งพิงกันในเชิงสถิติ

เนื่องจาก

$$E(XY) = (-2)(-2)(0.4) + (-2)(2)(0.1) + 0(-2)(0.4)$$
$$= 16 - 0.4 - 0.4 + 1.6 = 2.4$$
$$E(X) = (-2)(0.5) + (2)(0.5) = 0$$
$$E(X) = (-2)(0.5) + (2)(0.5) = 0$$
$$V(X) = (-2-0)^2(0.5) + (2-0)^2(0.5) = 4$$
$$V(Y) = (-2-0)^2(0.5) + (2-0)^2(0.5) = 4$$

ดังนั้น $Cov(x, y) = 2.4 - 0(0) = 2.4$

และ $\rho_{xy} = \dfrac{2.4}{\sqrt{4}\sqrt{4}} = 0.6$

ตัวอย่าง $f(x, y) = 2;\ 0 < x < y < 1$

ฟังก์ชั่นหนาแน่นน่าจะเป็นทางเดียวของ X และ

$$g(x) = \int_x^1 (2)\, dy = 2(1 - x);\ 0 < x < 1$$

$$h(y) = \int_0^y (2)\, dx = 2y;\ 0 < y < 1$$

ดังนั้น

$$E(X) = \int_0^1 x\, 2(1 - x)\, dx = 1/3$$

$$E(X^2) = \int_0^1 x^2\, 2(1 - x)\, dx = 1/6$$

3.9 การแจกแจงความน่าจะเป็นที่น่าสนใจ (Some Special Probability Distributions)

ตัวแปรเชิงสุ่มต่าง ๆ ที่กำหนดจากการทดลองหรือปรากฏการณ์เชิงสุ่มนั้นมีฟังก์ชั่นน่าจะเป็นในรูปฟอร์มมาตรฐานที่น่าสนใจ ฟังก์ชั่นน่าจะเป็นหรือการแจกแจงของตัวแปรเชิงสุ่มเหล่านั้นได้ใช้ประยุกต์ในเรื่องต่าง ๆ มากมาย การแจกแจงของตัวแปรเชิงสุ่มที่น่าสนใจมีดังนี้

3.9.1 การแจกแจงน่าจะเป็นของตัวแปรเชิงสุ่มไม่ต่อเนื่อง

(1) การแจกแจงยูนิฟอร์ม (Uniform Distribution) การทดลองเชิงสุ่มที่ให้ผลทดลองต่าง ๆ กัน แต่มีโอกาสที่เกิดขึ้นเท่า ๆ กัน (Mutually exclusive and Equaliy likely) นั้น ถ้ากำหนดตัวแปรเชิงสุ่ม X จากการทดลองแบบนี้ซึ่งมีค่าต่าง ๆ กัน n ค่า คือ $x_1, x_2, \ldots, x_n$ แล้วตัวแปรเชิงสุ่มจะมีการแจกแจงยูนิฟอร์ม ดังนี้

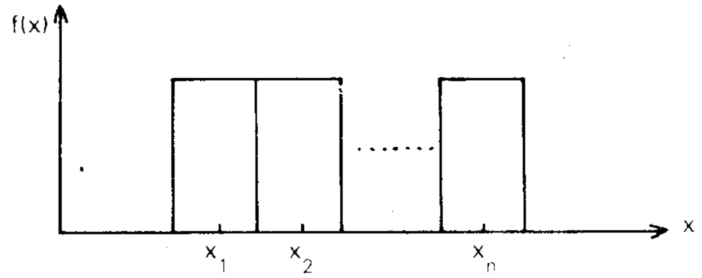
$$f(x) = 1/n ; x = x_1, x_2, \ldots, x_n$$

ค่าคาดหวังและความแปรปรวนจะเป็นดังนี้

$$E(X) = \Sigma X_i/n$$

$$V(x) = \Sigma \left\{X_i - E(X)\right\}^2/n$$

กราฟของการแจกแจงยูนิฟอร์มแสดงได้ดังนี้



เนื่องจากกราฟของการแจกแจงมีลักษณะเป็นแท่งสี่เหลี่ยมผืนผ้า จึงมีชื่อเรียกอีกอย่างว่า การแจกแจงสี่เหลี่ยม (Rectangular Distribution)

ตัวอย่าง โยนลูกเต๋า 1 ลูก ถ้าให้ X เป็นหน้าลูกเต๋าที่ขึ้น แล้ว X จะมีการแจกแจงยูนิฟอร์ม ดังนี้

$$f(x) = 1/6 ; x = 1, 2, \ldots, 6$$

(2) การแจกแจงเบอร์นุยยี (Bernoulli Distribution) ถ้าตัวแปรเชิงสุ่ม X กำหนดจากการทดลองเชิงสุ่ม (Random Trial) ที่มีกลุ่มผลทดลองประกอบด้วย 2 ผลทดลอง โดยเรียกผลทดลองแรกว่า "ความสำเร็จ (Success, S)" และผลทดลองหลังว่า "ไม่สำเร็จ (Failure, F)" , และถ้าความน่าจะเป็นของความสำเร็จเป็น P และความไม่สำเร็จเป็น Q = 1 - P ซึ่ง X มีค่าเป็น

x = 0 ถ้าผลทดลองเป็น F

x = 1 ถ้าผลทดลองเป็น S

แล้ว X จะมีการแจกแจงเบอร์นุยยี ดังนี้

$f(x) = P ; x = 1, 0 < P < 1$

$= Q ; x = 0, Q = 1 - P$

การทดลองเชิงสุ่มแบบนี้เป็นการทดลองที่ง่ายที่สุดในบรรดาการทดลองเชิงสุ่มทั้งหลาย ซึ่งมีชื่อว่าการทดลองแบบเบอร์นุยยี(Bernoulli's Trial) โดยตั้งตามนักคณิตศาสตร์ชื่อJaques Bernoulli (1654 - 1705)

ตัวแปรเชิงสุ่มนี้มีค่าคาดหวังและความแปรปรวน ดังนี้

$E(X) = P, V(X) = PQ$

ตัวอย่าง สุ่มคนงานหนึ่งจากคนงานของบริษัทที่มีคนงาน 60% เห็นด้วยกับแผนการของบริษัทเกี่ยวกับเงินเดือนใหม่

เมื่อให้ x มีค่าเป็น 1 ถ้าเขาเห็นด้วย และมีค่าเป็น 0 ถ้าเขาไม่เห็นด้วย แล้ว x จะมีการแจกแจงดังนี้

$f(x) = 0.60 ; x = 1$

$= 0.40; x = 0$

ค่าคาดหวังและความแปรปรวนจะเป็น

$E(X) = 0.60$

$V(X) = 0.60 (0.40) = 0.24$

การทดลองที่ให้ผลทดลองเป็นเหตุการณ์ E หรือไม่ก็ E เช่นสำเร็จและไม่สำเร็จ, ใช่และไม่ใช่, ชายและหญิง ของเสียและของดีเหล่านี้จะเกี่ยวข้องกับการแจกแจงเบอร์นุยยี

(3) **การแจกแจงทวินาม** (Binomial Distribution) ตัวแปรเชิงสุ่ม X ซึ่งเป็นจำนวนครั้งของความสำเร็จนั้นถ้ากำหนดจากการทดลองที่มีคุณสมบัติต่อไปนี้จะเรียกว่าตัวแปรเชิงสุ่มทวินาม

(1) การทดลองจะกระทำซ้ำ ๆ กัน จำนวนหนึ่งซึ่งจำกัดจำนวนครั้งและจะกระทำภายใต้สภาวการณ์เดียวกัน

(2) ผลทดลองที่เกิดขึ้นจากการทดลองแต่ละครั้งจะเป็นประเภทหนึ่งใน 2 ประเภท นั่นคือ สำเร็จ หรือไม่สำเร็จ นั่นเอง

(3) ความน่าจะเป็นของความสำเร็จ หรือ P(S) ในการทดลองครั้งหนึ่ง จะให้เป็น P ซึ่งจะคงที่ตลอดทุกการทดลอง

(4) ในการทดลองแต่ละครั้งจะเป็นอิสระกันหรือกว่าสั้น ๆ ได้ว่าตัวแปรเชิงสุ่ม X จำนวนครั้งของความสำเร็จอันเป็นผลจากการทำการทดลองแบบเบอร์นุยยีซ้ำ ๆ กัน n ครั้ง ภายใต้สภาวการณ์เดียวกัน และแต่ละครั้งเป็นอิสระกันนั้นจะเรียกว่า ตัวแปรเชิงสุ่มทวินาม โดยมีฟังก์ชัน ดังนี้

$$f(x) = \binom{n}{x} P^x (1-P)^{n-x} \quad x = 0, 1, 2, \ldots, n$$

ค่าคาดหวังและความแปรปรวนกำหนดไว้ดังนี้

$$E(X) = nP;\ V(X) = nPQ,\ Q = 1 - p$$

ตัวอย่าง ถ่านไฟฉายที่ผลิตจากเครื่องจักรจะใช้การไม่ได้ 10% ถ้าสุ่มถ่านไฟฉายที่ผลิตมา 3 ก้อน จงหาความน่าจะเป็นที่ถ่านนี้ไม่เสียเลย, เสีย 1 ก้อน, เสีย 2 ก้อน, เสียทั้งหมด และจงหาค่าคาดหวังและความแปรปรวนของจำนวนถ่านที่จะเสีย

ถ้า X เป็นจำนวนที่เสียแล้ว X จะมีค่าดังนี้ x = 0, 1, 2, 3, และมีการแจกแจงเป็น $f(x) = \binom{3}{x}(0.10)^x(1-0.10)^{3-x};\ x = 0, 1, 2, 3$

ดังนั้น $f(0) = \binom{3}{0}(0.10)^0(1-0.10)^{3-0} = (0.90)^3$

$=$

$$f(1) = \binom{3}{1}(0.10)^1\ (1-0.10)^{3-1} = 3(0.10)(0.90)^2$$

$$f(2) = \binom{3}{2}(0.10)^2(1-0.10)^{3-2} = 3(0.10)^2(1-0.10)$$

$$f(3) = \binom{3}{3}(0.10)^3(1-0.10)^{3-3} = (0.10)^3$$

และ $E(X) = 3(0.10) = 0.30$

$V(X) = 3(0.10)(0.90)$

(ก) **ฟังก์ชั่นแจกแจงสะสมแบบทวินาม**(Binomial Distribution Function) ฟังก์ชั่นแจกแจงทวินามจะให้ความน่าจะเป็นที่จะได้ความสำเร็จเป็นจำนวน X ครั้ง หรือน้อยกว่าจากการทดลอง n ครั้ง นั่นคือ $F(x) = P(X \leq x) = \sum_{s=0}^{x}\binom{n}{s} p^s(1-p)^{n-s}$

สำหรับความน่าจะเป็นที่จะได้ความสำเร็จเราอย่างน้อย x ครั้ง หรือ $P(X \geq x)$ กำหนดไว้ดังนี้ $P(X \geq x) = \sum_{s=x}^{n}\binom{n}{s} p^s(1-p)^{n-s}$

$$= 1 - P(X \leq x-1) = 1 - F(x-1)$$

ดังนั้น $P(X \leq x)$ เราจะเขียนใหม่ได้เป็น $P(X \leq x) = 1 - P(X \geq x+1)$

จากตัวอย่างของถ่านไฟฉายความน่าจะเป็นที่จะมีถ่านเสีย 2 ก้อน หรือน้อยกว่าจะหาได้

ดังนี้ $P(X \leq 2) = \sum_{s=0}^{2}\binom{3}{s}(.10)^s(1-.10)^{3-s} = f(0) + f(1) + f(2)$

หรือ $P(X \leq 2) = 1 - P(X \geq 3)$

$= \quad 1 - f(3) = 1 - (0.10)^3$

ถ้าจะหาความน่าจะเป็นที่มีถ่านเสียอย่างน้อย 1 ก้อน จะหาได้ดังนี้

$$P(X \geq 1) = \sum_{s=1}^{3} \binom{3}{s} (.10)^s (1-.10)^{3-s} = f(1) + f(2) + f(3)$$

หรือ $\quad P(X \geq 1) = 1 - P(X \leq 0) = 1 - f(0) = 1 - (0.90)^3$

การหาความน่าจะเป็นในแต่ละค่าของตัวแปรเชิงสุ่มทวินาม หรือความน่าจะเป็นสะสม ถ้าทราบค่า n และ P จะหาได้จากตารางการแจกแจงทวินามในท้ายหนังสือสถิติต่าง ๆ เช่น ถ้า n = 3, P = .10 จากตารางเราได้ f(0) = .0729, f(1) = .243, f(2) = .0270, f(3) = .0010 และ $P(X \geq 1) = .2710$, $P(X \geq 2) = .0280$, $P(X \geq 3) = .0010$ เป็นต้น

(ข) การแจกแจงทวินามยังมีความสัมพันธ์กับการแจกแบบปกติ และแบบเอฟอีกด้วย ซึ่งจะได้กล่าวต่อไป

(4) การแจกแจงพหุนาม (Multinomial Distribution) การแจกแจงพหุนามเป็นรูปทั่ว ๆ ไปของการแจกแจงทวินาม นั่นคือเกี่ยวข้องกับการทดลองที่ผลทดลองแบ่งเป็นประเภทมากกว่า 2 ประเภท และการทดลองนั้นมีคุณสมบัติดังนี้

(1) ผลทดลองของการทดลองแต่ละครั้งจะเป็นประเภทหนึ่งใน k ประเภท ($E_1, E_2, \ldots, E_k$)

(2) ความน่าจะเป็นของประเภทต่าง ๆ เหล่านี้ให้เท่ากับ $P_1, P_2, \ldots, P_k$ ตามลำดับ และความน่าจะเป็นเหล่านี้จะคงที่เสมอสำหรับการทดลองครั้งหนึ่ง ๆ

(3) การทดลองครั้งหนึ่ง ๆ จะเป็นอิสระกับการทดลองครั้งอื่น ๆ ทั้งหมด

(4) การทดลองต้องกระทำซ้ำ ๆ กันเป็นจำนวนหนึ่งที่จำกัดครั้งคือ n

ถ้าให้ $X_1, X_2, \ldots, X_k$ เป็นจำนวนครั้งที่ $E_1, E_2, \ldots, E_k$ เกิดขึ้นในการทดลอง n ครั้ง ซึ่ง $X_1, X_2, \ldots, X_k$ จะมีค่าเป็น 0, 1, 2, ..., n แล้วเราจะได้ว่า $X_1, X_2, \ldots, X_k$ จะเป็นตัวแปรเชิงสุ่มพหุนามที่มีการแจกแจงเป็น

$$f(x_1, x_2, \ldots, x_k) = \frac{n!}{(x_1!)(x_2!)\ldots(x_k!)} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

ในเมื่อ $x_i = 0, 1, 2, \ldots, n \ (i = 1, 2, \ldots, k)$; $\Sigma x_i = n$; $\Sigma p_i = 1$

ในการแจกแจงพหุนามนั้นจำนวนครั้งที่เกิดขึ้นของ E_i ใด ๆ ($i = 0, 1, 2, \ldots, k$)

ในการทดลองซ้ำ ๆ กัน n ครั้งที่เป็นอิสระกันนั้นอาจจะถือว่าเป็น**ตัวแปรเชิงสุ่มทวินาม**ตัวหนึ่ง (X_i) ได้โดยจะมีคุณสมบัติดังนี้

(1) $E(X_i) = nP$

(2) $V(X_i) = nP_i(1-P_i); i = 1, 2, \ldots, k$

(3) $Cov(X_i, X_j) = -nP_iP_j$

(4) $\rho_{ij} = \sqrt{P_iP_j/(1-P_i)(1-P_j)}$

(5) $\sum_{i=1}^{h} (X_i - nP_i)^2/nP_i$ จะมีการแจกแจงไคสแควร์โดยประมาณด้วยองศาความเป็นอิสระ $k-1$ ในทางปฏิบัติ nP จะไม่น้อยมากจนเกินไป นั่นคือ $nP > 5$

ตัวอย่าง บริษัทโภคผลมีคนงานอยู่จำนวนมาก ครึ่งหนึ่งของคนทำงานอยู่ฝ่ายผลิต หนึ่งในสามทำหน้าที่เป็นพนักงานขาย และที่เหลืออยู่ฝ่ายบริหารบุคคล บริษัทต้องการตั้งคณะกรรมการคณะหนึ่ง ซึ่งเลือกจากคนงานทั้งหมด

จงหาความน่าจะเป็นที่คณะกรรมการชุดนี้จะทำงานอยู่ฝ่ายผลิตสามคน พนักงานขายสองคน และฝ่ายบริหารบุคคลหนึ่งคน

ให้ E_1 คนงานฝ่ายผลิต $P(E_1) = p_1 = 1/2$

E_2 : คนงานฝ่ายพนักงานขาย $P(E_2) = p_2 = 1/3$

E_3 : คนงานฝ่ายบริหารบุคคล $P(E_3) = p_3 = 1/6$

และให้ X_1, X_2 และ X_3 เป็นจำนวนคนงานฝ่ายการผลิต, พนักงานขายและบริหารบุคคลตามลำดับ แล้ว X_1, X_2 และ X_3 มีการแจกแจงดังนี้

$$f(x_1, x_2, x_3) = \frac{6!}{x_1!x_2!x_3!} (1/2)^{x_1} (1/3)^{x_2} (1/6)^{x_3}$$

ในเมื่อ $x_i = 0, 1, 2, \ldots, 6; \sum p_i = 1, \sum x_i = 6$

ดังนั้น ความน่าจะเป็นที่ต้องการคือ $P(X_1 = 3, X_2 = 2, X_3 = 1) = f(3,2,1)$

$$= \frac{6!}{3!\,2!\,1!} (1/2)^3 (1/3)^2 (1/6) = 5/36$$

สำหรับค่าคาดหวังและความแปรปรวนของ X_1, X_2 และ X_3 หาได้ดังนี้

$$E(X_1) = nP_1 = 6(1/2) = 3, \quad V(X_1) = nP_1(1-P_1) = 6(1/2)(1/2) = 1.5$$
$$E(X_2) = nP_2 = 6(1/3) = 2, \quad V(X_2) = nP_2(1-P_2) = 6(1/3)(1/3) = 4/3$$
$$E(X_3) = nP_3 = 6(1/6) = 1, \quad V(X_3) = nP_3(1-P_3) = 6(1/6)(5/6) = 5/6$$

ตัวอย่าง โรงงานจะทดสอบความเรียบร้อยของผลิตภัณฑ์ที่ผลิตออกมาหนึ่งหน่วยต่อทุก 100 หน่วย เป็นที่ทราบกันว่า ความน่าจะเป็นของผลิตภัณฑ์ที่ผลิตออกจากโรงงานจะไม่เรียบร้อยเท่ากับ 0.2

ถ้าหยิบของที่ผลิตออกมา 5 หน่วย จงหาความน่าจะเป็นของผลิตภัณฑ์ที่เรียบร้อยแต่ไม่ถูกทดสอบ, ที่เรียบร้อยและถูกทดสอบ, ที่ไม่เรียบร้อยแต่ถูกทดสอบ, และที่ไม่เรียบร้อยและ

ไม่ถูกทดสอบ มีจำนวนเป็น 3, 1, 1 และ 0 ตามลำดับ

ผลิตภัณฑ์ที่ผลิตออกมาจะมีกลุ่มผลทดลองดังตารางต่อไปนี้

	ความเรียบร้อย	เรียบร้อย (D_1)	ไม่เรียบร้อย (D_2)
การทดสอบ	ถูกทดสอบ (I_1)	I_1D_1	I_1D_2
	ไม่ถูกทดสอบ (I_2)	I_2D_1	I_2D_2

ดังนั้น $P(I_1D_1) = .01(1-0.2) = .01(0.8)$

$P(I_1D_2) = .01(0.2)$

$P(I_2D_1) = (1-0.01)(1-0.2) = .99(0.8)$

$P(I_2D_2) = (1-0.01)(0.2) = 0.99(0.2)$

ถ้าให้ X_1, X_2, X_3 และ X_4 เป็นจำนวนผลิตภัณฑ์ที่เป็น $(I_1D_1), (I_1D_2)$, (I_2D_1) และ (I_2D_2) แล้วความน่าจะเป็นที่ต้องการคือ

$P(X_1 = 1, X_2 = 1, X_3 = 3, X_4 = 0) = f(1, 1, 3, 0)$

$$= \frac{5!}{1!\,1!\,3!\,0!}\{0.01(.8)\}^1\{.01(.2)\}^1\{.99(8)\}^3\{.99(.2)\}^0$$

$$= 20(.01)^2(.99)^2(.99)^3(.8)^4(.2) = .00016$$

(5) **การแจกแจงทวินามนิเสธ** (Negative Binomial Distribution) ในการทำการทดลองแบบเบอร์นุยยีซ้ำ ๆ กันจนกว่าจะประสบความสำเร็จ k ครั้ง แล้วจึงเลิกทำนั้น ถ้าให้ X เป็นตัวแปรเชิงสุ่มที่แทนจำนวนครั้งที่ต้องการทดลองจนกว่าจะสำเร็จ k ครั้ง แล้ว X จะมีค่าเป็น k, k + 1,.... และมีการแจกแจง ดังนี้

$$f(x) = \binom{x-1}{k-1} P^k (1-P)^{x-k}; \quad x = k, k+1, \ldots$$

หรือ

$$f(y) = \binom{y+k-1}{k-1} P^k (1-P)^y; \quad y = 0, 1, 2, \ldots \qquad y = x-k$$

ค่าคาดหวังและความแปรปรวนกำหนดไว้ดังนี้

$E(X) = k/p$, $V(X) = k(1-p)/p^2$

การแจกแจงนี้เรียกอย่างหนึ่งว่า การแจกแจงปาสคาล (Pascal Distribution)

ตัวอย่าง พนักงานขายของในบริษัท กขค ออกไปเสนอขายผลิตภัณฑ์ให้แก่ลูกค้าเป็นทราบกันว่าลูกค้าประมาณ 40 % จะซื้อของตามที่เสนอ

ถ้าพนักงานขายคนหนึ่งต้องการขายของให้ได้ 3 ชิ้นก่อนสิ้นปี เพื่อจะได้ติดอันดับพนักงานดีเด่นประจำปี

จงหาโอกาสที่เขาจะได้เป็นพนักงานดีเด่น ถ้าภายในกำหนดนั้นเขาเสนอขายของให้แก่

ลูกค้าเพียง 5 รายเท่านั้น

เรามี $k = 3, P = 0.4, (1-P) = 0.6$

ดังนั้น $P(X=5) = \binom{5-1}{3-1}(.4)^3(.6)^{5-3} = .13824$

สำหรับ E(X) และ V(X) คำนวณได้ดังนี้

$$E(X) = 3/0.4 = 7.5$$

$$V(X) = 3.(.6)/(.4)^2 = 11.25$$

(ก) การแจกแจงปาสคาลมีความสัมพันธ์กับการแจกแจงทวินามดังนี้ ถ้า X มีการแจกแจงทวินามซึ่งมี**พารา**มิเตอร์ n และ P

ส่วน Y มีการแจกแจงปาสคาลที่มีพารามิเตอร์ k และ p แล้ว X และ Y มีความสัมพันธ์กันดังนี้

$$P(Y \le n) = P(X \ge k)$$

หรือ
$$\sum_{y=k}^{n}\binom{y-1}{k-1}p^k q^{y-k} = \sum_{x=k}^{n}\binom{n}{x}p^x q^{n-x}$$

ดังนั้น การคำนวณความนาจะเป็นของการแจกแจงปาสคาลจะคำนวณได้จากการแจกแจงทวินามได้อีกด้วย ดังเช่นตัวอย่างต่อไปนี้

ตัวอย่าง พนักงานขายเสนอขายของแก่ลูกค้าจำนวนหนึ่ง เป็นที่ทราบว่า ลูกค้าจะสั่งของเพียง 20% ถ้าเขาต้องการให้ลูกค้าสั่งถึง 5 ราย แล้วโอกาสที่เขาจะต้องติดต่อลูกค้าไม่เกิน 25 รายเป็นเท่าใด

ให้ Y เป็นจำนวนลูกค้าที่พนักงานเสนอขาย และมีลูกค้าสั่งของ k = 5 รายดังนั้น Y จะมีค่าเป็น 5, 6, 7,....และมีการแจกแจงดังนี้

$$f(y) = \binom{y-1}{5-1}(.2)^5\ (.8)^{y-5},\ y = 5,6,\ldots$$

$$P(Y \le 25) = \sum_{y=5}^{25} f(y) = \sum_{x=5}^{25}\binom{25}{x}(.2)^x(.8)^{25-x}$$

$$= 1-\sum_{x=0}^{4}\binom{25}{x}(.2)^x\ (.8)^{25-x}$$

$$= 1-.42067 = .57933$$

(ข) การแจกแจงปาสคาลสามารถประมาณได้ด้วยการแจกแจงปกติ ซึ่งจะได้กล่าวต่อไป

(6) การแจกแจงเรขาคณิต (Geometric Distribution) สำหรับตัวแปรเชิงสุ่ม X ที่เป็นแบบทวินามนิเสธ ซึ่งเกี่ยวข้องกับการทำ การทดลองแบบเบอร์นุยยีที่ทำซ้ำ ๆ จนประสบความสำเร็จครั้งแรกแล้วเลิก นั่นคือ k = 1 แล้ว X จะมีการแจกแจงเรขาคณิตฟังก์ชั่นต่อไปนี้

$f(x) = P(1-P)^{x-1}$; x = 1, 2, 3,....

ค่าคาดหวังและความแปรปรวนหาได้ดังนี้

$E(X) = 1/p$, $V(X) = (1-p)/p^2$

ตัวอย่าง สิงห์ตั้งใจซื้อลอตเตอรี่งวดละใบจะถูกรางวัลที่ 1 จงหาโอกาสที่เขาจะถูกรางวัลที่ 1 ถ้าเขาซื้องวดที่ 100 และถ้าให้เขาถูกรางวัลที่ 1 สักครั้งหนึ่ง เขาควรจะซื้อลอตเตอรี่สักกี่ครั้ง?

$P = 1/1{,}000{,}000$ (ถ้าสมมติสลากออกงวดละล้านใบ)

ดังนั้น $P(X = 100) = .000001\ (.999999)^{100-1} \cong 0$ และเขาควรซื้อลอตเตอรี่เท่ากับค่าคาดหวังของจำนวนครั้งที่ซื้อ คือ E(X)

$E(X) = 1/P = 1/(.000001) = 1{,}000{,}000$ ครั้ง

ตัวอย่าง ค่าใช้จ่ายทำการทดลองครั้งหนึ่งเป็นเงิน 20,000 บาท แต่ถ้าการทดลองประสบความล้มเหลวต้องเพิ่มค่าใช้จ่ายอีก 6,000 บาท เพราะต้องเปลี่ยนแปลงบางส่วนก่อนที่จะทำการทดลองครั้งต่อไป สำหรับความน่าจะเป็นที่ประสบความสำเร็จในการทดลองครั้งหนึ่ง ๆ เป็น 0.2 และการทดลองแต่ละครั้งเป็นอิสระกัน

ถ้าทำการทดลองจนกว่าจะประสบความสำเร็จครั้งหนึ่ง จะต้องเสียค่าใช้จ่ายเฉลี่ยในกระบวนการทั้งหมดเท่าใด?

ให้ C เป็นค่าใช้จ่าย และ X เป็นจำนวนครั้งที่จะต้องทำจนกว่าจะประสบความสำเร็จ เราจะได้ว่า

$$C = 20000(X)+6000\,(X-1)$$

$$= 26000(X) - 6000$$

ดังนั้น $E(C) = 26000\,E(X) - 6000$

$$= 26000\,(1/0.2)-6000 = 124000$$

(7) **การแจกแจงไฮเปอร์จีโอเมตริค** (Hypergeometric Distribution) ในการแจกแจงทวินามนั้นเกี่ยวข้องกับการทดลองเชิงสุ่มซ้ำ ๆ กันที่เป็นอิสระ แต่ถ้าไม่เป็นอิสระต่อกันและความน่าจะเป็นที่จะประสบความสำเร็จครั้งใด ๆ ไม่คงที่ แล้วการทดลองแบบนี้จะเกี่ยวข้องกับการแจกแจงไฮเปอร์จีโอเมตริค

โดยทั่วไปการแจกแจงนี้เกี่ยวข้องกับตัวอย่างสุ่มขนาด n หน่วย ที่ดึงมาจากประชากรจำกัดที่มีขนาด N และประชากรแยกออกได้ 2 พวก คือพวกแรกเรียกว่าระบบความสำเร็จ (Success) ซึ่งมี m หน่วย พวกที่สองเรียกว่าไม่ประสบความสำเร็จ (Failure) จะมี (N-m) หน่วย

ถ้า X เป็นจำนวนที่ประสบความสำเร็จ ในตัวอย่างสุ่ม n หน่วย ที่เลือกมาแล้ว X จะมีค่าเป็น $\max\{0, m-(N-m)\} \leq x \leq \min(n,m)$ และมีการแจกแจงดังนี้

$$f(x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{n}{x}} \; ; \max\{0, n-(N-m)\} \leq x \leq \min(n,m)$$

ในเมื่อ min (n,m) **คือจำนวนที่เลือกเอาตัวน้อยที่สุดระหว่าง** n กับ m **เช่น** min (5, 10) = 5 , min (10, 3) = 3 **เป็นต้น และ** max { o, n − (n, m)} **คือจำนวนที่เลือกเอาตัวมากที่สุดระหว่าง** 0 กับ n − (N − m) **เช่น** max (O, 5) = 5 , max (O,-2) = 0 **เป็นต้น**

สำหรับค่าคาดหวังและความแปรปรวนเป็นดังนี้

$E(X) = n(m/N)$

$V(X) = n(m/N)\,(1-m/N)\,\frac{N-n}{N-1}$

จะเห็นได้ว่าค่าคาดหวังของการแจกแจงไฮเปอร์จีโอเมตริคจะเท่ากับของแบบทวินาม (ในเมื่อ m/N = p นั่นเอง) แต่ความแปรปรวนต่างกันด้วยตัวคูณ $\frac{N-n}{N-1}$ ถ้า N มีมากพอ $[N \to \infty]$ แล้วตัวคูณ $\frac{N-n}{N-1}$ จะเท่ากับ 1 ซึ่งจะได้ความแปรปรวนเท่ากันกับของการแจกแจงทวินาม

ในกรณีที่ประชากรมีขนาดโตพอ $(N \to \infty)$ การแจกแจงไฮเปอร์จีโอเมตริคจะมีแนวโน้มเท่ากับการแจกแจงทวินาม นั่นคือ

$$\lim_{N \to \infty} \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}} \longrightarrow \binom{n}{x} P^x (1-P)^{n-x}; P = m/N$$

การประยุกต์ของการแจกแจงไฮเปอร์จีโอเมตริค การแจกแจงนี้ใช้สำหรับประมาณหน่วยทั้งหมดของประชากรที่เราไม่สามารถแจงนับทั้งหมดได้ เช่น สัตว์ในป่า ปลาในน้ำ และยังสามารถประมาณจำนวนในลักษณะหนึ่งของประชากรได้อีก เช่น ประมาณจำนวนรถที่เสียจากผลิตมาทั้งหมด

(ก) การประมาณจำนวนประชากร ลองพิจารณาตัวแบบการสุ่มตัวอย่างทางตรง (Direct Sampling Model) ต่อไปนี้ ในการประมาณจำนวนปลาในทะเลสาปจะมีขบวนการในการแก้ปัญหา ดังนี้

จับปลามาจำนวนหนึ่งจากทะเลสาป สมมติว่า m ตัว แล้วทำเครื่องหมายไว้ และปล่อยปลาลงไปในทะเลสาปตามเดิม

ต่อมาจับปลามาอีกจำนวนหนึ่ง สมมติให้เป็น n ตัว แล้วนับจำนวนปลาที่มีเครื่องหมาย

ถ้า X เป็นจำนวนปลาที่มีเครื่องหมายของการจับปลาครั้งที่สอง และ N เป็นจำนวนปลาทั้งหมดในทะเลสาป ดังนั้นการแจกแจงของ X จะเป็น

$$f(x) = \frac{\binom{m}{n}\binom{N-m}{n-x}}{\binom{N}{n}}$$

สำหรับค่าของ N ที่จะประมาณ ซึ่งจะให้เป็น $\hat{N}$ นั้น เรามีวิธีการประมาณด้วยวิธีการหนึ่ง คือการประมาณแบบน่าจะเป็นมากที่สุด (Maximum Likelihood Estimate) และจะได้ $\hat{N}$ ดังนี้

$$\hat{N} = [nm/x]$$

$\hat{N}$ เป็นจำนวนเต็มบวกที่มากที่สุดจากความสัมพันธ์ $\hat{N} \leq \frac{nm}{x}$ เช่น ถ้าจับปลาครั้งแรกมา 2,000 ตัว ทำเครื่องหมายแล้วปล่อยไป ตอนหลังจับปลามาอีก 1,000 ตัว นับจำนวนปลาที่มีเครื่องหมายได้ 20 ตัว แล้วปลาในทะเลสาปจะมีประมาณ

$$\hat{N} = [\frac{nm}{x}] = [1000(2000)/20] = 100000 \text{ ตัว}$$

(ข) ในการประมาณจำนวนของลักษณะหนึ่งที่สนใจ เช่น ในการประมาณจำนวนของเสียในการผลิตครั้งหนึ่ง เราทำได้ดังนี้ ถ้า N เป็นขนาดของการผลิตครั้งหนึ่งที่ทราบ n เป็นขนาดตัวอย่างที่สุ่มของมา x เป็นจำนวนของเสียในตัวอย่าง และ m เป็นจำนวนของเสียที่เราต้องการทราบค่า ดังนั้น โอกาสที่ของเสีย x ในตัวอย่างขนาด n ที่สุ่มจากการผลิตครั้งหนึ่งซึ่งมีขนาด N โดยมีของเสีย m จะเท่ากับ

$$\binom{m}{x}\binom{N-m}{n-x} \Big/ \binom{N}{n}$$

ในการประมาณค่า m คือ $\hat{m}$ ทำได้แบบเดียวกับ $\hat{N}$ ซึ่งจะได้ดังนี้

$$\hat{m} = [(N+1)x/n]$$

ในเมื่อ $\hat{m}$ เป็นจำนวนเต็มบวกที่ได้จากความสัมพันธ์ $\hat{m} \leq (N+1)x/n$

ตัวอย่าง รถส่งมาจากญี่ปุ่น 1,000 คัน ผู้ตรวจสอบรถสุ่มรถมา 25 คัน จากการตรวจสอบพบว่ามี 1 คัน ที่ระบบไฟใช้ไม่ได้ ดังนั้นจำนวนรถที่ระบบไฟใช้ไม่ได้มีประมาณ $\hat{m}$

$$\hat{m} = \left[\frac{(N+1)x}{n}\right] = \frac{(1000+1)(1)}{25} = [40.04]$$

ดังนั้น $\hat{m} = 40$ คัน

(8) **การแจกแจงไฮเปอร์จีโอเมตริกทั่วไป** (Generalized Hypergeometric Distribution) การแจกแจงแบบนี้เป็นผลมาจากการสุ่มตัวอย่างขนาด n จากประชากรที่มีขนาด N โดยที่ประชากรแบ่งออกได้เป็น k พวก $(E_1, E_2, \ldots, E_k)$ แต่ละพวกมีจำนวน $m_1 m_2, \ldots, m_k$ ตามลำดับ ถ้าให้ $x_1, x_2, \ldots, x_k$ เป็นจำนวนครั้งที่เหตุการณ์ $E_1, E_2, \ldots, E_k$ เกิดขึ้นแล้ว $X_1, X_2, \ldots, X_k$ จะมีการแจกแจงไฮเปอร์จีโอเมตริกทั่วไปดังนี้

$$f(x_1, x_2, \ldots, x_k) = \binom{m_1}{x_1}\binom{m_2}{x_2}\cdots\binom{m_k}{x_k} \Big/ \binom{N}{n}$$

ในเมื่อ $x_i = 0, 1, 2, \ldots, n$ $(i = 1,2,\ldots,k)$, $\Sigma x_i = n, \Sigma m_i = N$

ตัวอย่าง โรงงานทอผ้ามีคนงานอยู่ 40 คน ทำหน้าที่คุมเครื่องจักร 2 คน ฝ่ายบริหาร 3 คน พนักงานขาย 5 คน นอกนั้นทำหน้าที่ทอ พอสิ้นปีประธานกรรมการบริษัทให้รางวัลแก่คนงาน 5 รางวัล โดยการจับฉลากที่มีจำนวนฉลากเท่ากับคนงาน จงหาโอกาสมีจำนวนคนงานที่ทำหน้าที่การทอ, พนักงานขาย, ฝ่ายบริหาร, และคุมเครื่องจักรได้รับรางวัลเป็น 2,1,1,1 คน ตามลำดับ

$$P(X_1=2, X_2=1, X_3=1, X_4=1) = f(2,1,1,1) = \binom{40}{2}\binom{5}{1}\binom{3}{1}\binom{2}{1}/\binom{50}{5}$$

(9) การแจกแจงไฮเปอร์จีโอเมตริคนิเสธ (Negative Hypergeometric Distribution) การแจกแจงแบบนี้เกี่ยวข้องกับตัวอย่างสุ่มที่ดึงมาจากประชากรที่มีขนาด N หน่วย จำนวน N หน่วยนี้แบ่งได้เป็น 2 พวก พวกหนึ่งเรียกสำเร็จ (Success) มีจำนวน m อีกพวกเรียกไม่สำเร็จ (Failure) มีจำนวน N-m ถ้าจะดึงตัวอย่างสุ่มให้ได้สำเร็จถึง k ครั้ง จำนวนครั้งหรือขนาดของตัวอย่างสุ่มจะเป็นตัวแปรเชิงสุ่ม และถ้าให้เป็น X แล้วมันจะมีค่าเป็น k, k, +1, k+2..... และมีการแจกแจงไฮเปอร์จีโอเมตริคนิเสธ ดังนี้

$$f(x) = \frac{\binom{m}{k-1}\binom{N-m}{x-k}}{\binom{N}{x-1}} \; \frac{m-k+1}{N-x+1}$$

$$= \frac{\binom{m}{k-1}\binom{N-m}{x-k}(x-1)!\;(m-k+1)}{(N)_x}$$

ในเมื่อ $x = k, k+1, \ldots$ และ $(N)_x = N_{P_x} = N!/(N-x)!$

การแจกแจงแบบนี้นำไปประยุกต์ในการประมาณจำนวนสัตว์ (Zoolgical Sample Census) ได้เช่นเดียวกับการแจกแจงไฮเปอร์จีโอเมตริค ดังตัวแบบการสุ่มตัวอย่างทางอ้อม (Inverse Sampling Model) ต่อไปนี้

ปัญหาที่พิจารณาก็เช่นเดียวกับที่กล่าวมาในตัวแบบการสุ่มตัวอย่างทางตรง วิธีการดำเนินไปตามขั้นดังนี้

- จับปลามาจำนวนหนึ่ง ทำเครื่องหมายและปล่อยปลาลงไปตามเดิม
- จับปลาอีกโดยจับให้ได้ปลาที่มีเครื่องหมายตามจำนวนที่ต้องการ

ถ้าให้ N เป็นจำนวนปลาในทะเลสาปที่ไม่ทราบค่า m จำนวนปลาที่ทำเครื่องหมายไว้ k จำนวนปลาที่มีเครื่องหมายซึ่งตั้งใจจะจับให้ได้ในครั้งที่สอง และ x เป็นจำนวนปลาที่ต้องจับจนกว่าจะได้ปลาที่มีเครื่องหมาย k ตัว แล้ว x จะเป็นตัวแปร เชิงสุ่มไฮเปอร์จีโอเมตริคนิเสธ ที่มีการแจกแจงดังนี้

$$f(x) = \frac{\binom{m}{k-1}\binom{N-m}{x-k}}{\binom{N}{x-1}} \quad \frac{m-k+1}{N-x+1} \qquad x=k,\ k+1,\ \ldots$$

ในทำนองเดียวกันกับวิธีแรกเราจะประมาณค่าของ N คือ $\hat{N}$ ได้ดังนี้

$$\hat{N} = (xm/k)$$

ตัวอย่าง ถ้าจับปลาครั้งแรกได้ 250 ตัว ทำเครื่องหมายและปล่อยลงไป จับปลาครั้งที่สองจนได้ปลาที่มีเครื่องหมาย 50 ตัว ซึ่งต้องทำการจับหลาตัวที่ 5000 แล้ว เราจะประมาณค่าของ N ได้ดังนี้

$$\hat{N} = (xm/k) = (250\ (5000)/50) = 25000$$

(10) **การแจกแจงปัวซอง** (Poisson Distribution) ตัวแปรเชิงสุ่ม x ซึ่งแทนจำนวนครั้งที่เกิดขึ้นของเหตุการณ์ในช่วง (interval) หรือเขต (Region) ที่กำหนดให้นั้นจะมีค่าเป็น 0, 1, 2,..... และมีการแจกแจงปังซองดังนี้

$$f(x) = \frac{e^{-\alpha}\alpha^{x}}{x!}\ ;\ x = 0, 1, 2, \ldots$$

ในเมื่อ α เป็นส่วนเฉลี่ยของจำนวนครั้งที่เกิดขึ้นของเหตุการณ์ที่สนใจภายในช่วงที่กำหนดให้ และ $e = 2.71828\ldots$

ตัวแปรเชิงสุ่มปัวซองนี้เกิดขึ้นแพร่หลายในธรรมชาติ เช่น จำนวนอุบัติเหตุที่เกิดขึ้นในคาบเวลาหนึ่ง จำนวนลูกค้าที่เข้ามาติดต่อกับบริษัทในช่วงเวลาหนึ่ง จำนวนยานพาหนะที่เข้าเทียบท่าเรือ ท่ารถบัส ท่าอากาศยาน เป็นต้น

ค่าคาดหวังและความแปรปรวนของตัวแปรเชิงสุ่มปัวซองนี้จะเท่ากันนั่นคือ

$$E(X) = \alpha,\ V(X) = \alpha$$

ถ้า α มีค่าน้อยแล้วกราฟของการแจกแจงปัวซองจะเบ้ทางขวามาก แต่ถ้าค่ายิ่งมากกราฟจะยิ่งเข้าสู่สมมาตรมากขึ้น

ตัวอย่าง บริษัทได้ทดสอบยางรถยนต์โดยขับรถไปตามแหล่งทุรกันดาร และได้พบว่าสาเหตุภายนอกที่ทำให้ยางแบบนั้นจะเกิดขึ้นเฉลี่ย 1 ครั้งต่อ 300 ไมล์ สมมุติจำนวนครั้งที่ยางแบบเกิดขึ้นมีการแจกแจงปัวซอง จงหาความน่าจะเป็นที่เดินทางไป 1800 ไมล์ แล้วไม่เกิดยางแบนขึ้นอันเนื่องจากสาเหตุภายนอก ?

ให้ x เป็นจำนวนครั้งที่ยางแบนเนื่องจากสาเหตุภายนอก และ α เป็นจำนวนเฉลี่ยที่ยางแบนเกิดขึ้นในระยะทาง 1800 ไมล์ นั่นคือ

$$\alpha = \frac{1}{3000}(1800) = 0.6$$

ดังนั้น $f(x) = e^{-0.6}(0.6)^x/x!$; $x = 0, 1, 2, \ldots$

นั่นคือ $P(x = 0) = f(0) = e^{-0.6}(0.6)^0/0! = 0.5488$

ถ้าทดสอบในระยะทาง 3600 ไมล์ จงหาความน่าจะเป็นที่ยางแบนครั้งหนึ่งอันเนื่องมาจากสาเหตุภายนอก

ดังนั้น $f(x) = e^{-1.2}(1.2)^x/x!$; $x = 0, 1, 2, \ldots$

ในเมื่อ $\alpha = (\frac{1}{3000})(3600) = 1.2$

นั่นคือ $P(X = 1) = f(1) = e^{-1.2}(1.2)^1/1! = 0.3614$

(ก) การประมาณการแจกแจงทวินามด้วยการแจกแจงปัวซอง ในทางปฏิบัติจะคำนวณความน่าจะเป็นจากการแจกแจงทวินามที่มี n โต ๆ ได้ยาก

นักคณิตศาสตร์จึงหาวิธีที่ไม่ยากนักและวิธีหนึ่งที่ใช้กันคือ ใช้การแจกแจงปัวซอง ซึ่งสมัยแรกรู้จักกันในนามของ "สูตรคำนวณความน่าจะเป็นที่จะประสบความสำเร็จ x ครั้ง เมื่อทำการทดลองแบบเบอร์นยยีซ้ำหลาย ๆ ครั้ง Simeon Denis Poisson (1781 - 1840) ได้พิสูจน์ว่า

$$\lim_{\substack{n \to \infty \\ p \to 0}} \binom{n}{x} P^x (1-P)^{n-x} \longrightarrow \frac{e^{-nP}(nP)^x}{x!}$$

ดังนั้น การแจกแจงปัวซองนอกจากจะใช้อธิบายปรากฏการณ์ธรรมชาติ แล้วมันยังจะใช้เป็นขีดจำกัดของการแจกแจงทวินามอีกด้วย ในเมื่อ n โต และ P น้อย

โดยทั่วไปการประมาณจะทำได้ถ้า $P \leq 0.05$ และ $n \geq 100$ อย่างไรก็ตามการประมาณจะเที่ยงตรงขึ้น ถ้า P น้อยมาก ๆ

ตัวอย่าง จากการศึกษาถึงคนติดสุราพบว่าชายไทย โดยเฉลี่ย 1 คน ต่อ 1000 คน จะติดสุรา ถ้าสุ่มคนมา 8000 คน โอกาสที่จะมีคนติดสุราน้อยกว่า 7 คน เป็นเท่าใด ?

$$n = 8000,\ P = 0.001 \quad \alpha = 8000(.001) = 8$$

$$P(x < 7) = \sum_{x=0}^{6} \binom{8000}{x} (.001)^x (.999)^{8000-x}$$

$$\cong \sum_{x=0}^{6} e^{-8}\, 8^x/x! = 0.3134$$

(ข) การแจกแจงปัวซองนี้สามารถประมาณได้ด้วยการแจกแจงปกติ ซึ่งจะได้กล่าวต่อไป

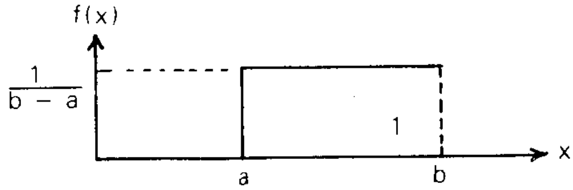
(ค) การแจกแจงปัวซองนี้มีความสัมพันธ์กับการแจกแจงไคสแควร์ด้วย ซึ่งจะได้กล่าวต่อไป

3.9.2 การแจกแจงน่าจะเป็นของตัวแปรเชิงสุ่มต่อเนื่อง

(1) **การแจกแจงยูนิฟอร์ม** (Uniform distribution) ตัวแปรงเชิงสุ่ม X ที่มีค่าอยู่ในช่วงของเลขจำนวนจริงสองจำนวน a และ b โดยที่ $a < b$ และ ถ้าฟังก์ชั่นหนาแน่นของ X นี้คงที่ตลอดช่วงจาก a ถึง b แล้วตัวแปรเชิงสุ่ม x จะมีการแจกแจงแบบยูนิฟอร์ม นั่นคือ

$$f(x) = \frac{1}{b-a} ; \; a < x < b$$

กราฟของการแจกแจงนี้เป็นดังนี้



สำหรับค่าคาดหวังและความแปรปรวนหาได้เป็นดังนี้

$E(x) = (b + a)/2, \; V(x) = (b - a)^2/12$

ตัวอย่าง รถเมล์วิ่งผ่านหน้ามหาวิทยาลัยทุก ๆ ครึ่งชั่วโมงในระหว่างเวลา 21.00 น. ถึง 24.00 น. โอกาสที่นักศึกษาคนหนึ่งเดินมาถึงป้ายรถเมล์หน้ามหาวิทยาลัยระหว่างคาบเวลานี้จะต้องรอรถอย่างน้อย 20 นาที เป็นเท่าใด ?

ให้ตัวแปรเชิงสุ่ม T เป็น เวลาที่รถคันต่อไปจะมาถึงป้าย ซึ่งมีการแจกแจงยูนิฟอร์มในช่วง $0 < t < 30$

โอกาสที่เขาจะต้องรออย่างน้อย 20 นาที คือ $P(T \geq 20)$

$$P(x \geq 20) = \int_{20}^{30} \left(\frac{1}{30-0}\right) dt = \frac{1}{3}$$

ค่าคาดหวังของเวลาที่จะต้องรอ $E(T) = (30 + 0)/2 = 15$ นาที

และความแปรปรวน $V(T) = (30 - 0)^2 / 12 = 75$ นาที

(2) การแจกแจงปกติ (Normal Distribution) การแจกแจงปกติดูเหมือนจะเป็นการแจกแจงที่สำคัญที่สุด เพราะว่าการแจกแจงของประชากรต่าง ๆ สามารถประมาณได้ด้วยการแจกแจงปกติโดยตัวแปลง (Transformations) ง่าย ๆ

ตัวแปรเชิงสุ่ม X จะเรียกว่ามีการแจกจงปกติ ถ้าฟังก์ชั่นหนาแน่นน่าจะเป็นของ X กำหนดไว้ดังนี้

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left\{\frac{x-\mu}{\sigma}\right\}^2}, \quad -\infty < x < \infty$$

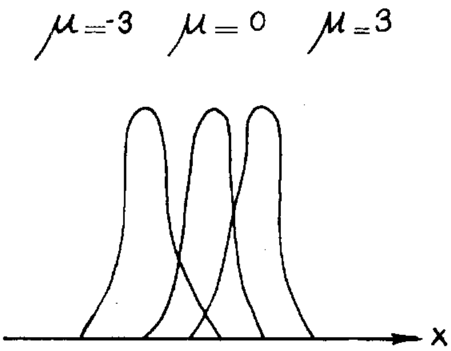
ในเมื่อ $\pi = 3.14159....$, μ และ σ เป็นพารามิเตอร์ โดยที่ $-\infty < \mu < \infty$ และ $\sigma > 0$

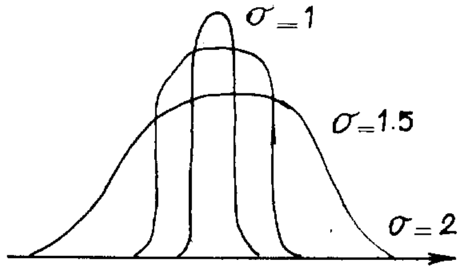
สำหรับค่าคาดหวังและความแปรปรวน คือ

$E(x) = \mu$ และ $V(x) = \sigma^2$

ถ้าค่า μ เปลี่ยนแปลงจะทำให้โค้งหรือกราฟของฟังก์ชั่นหนาแน่นปกติ ย้ายไปทางซ้ายหรือทางขวาได้แต่รูปร่างของโค้งไม่เปลี่ยนแปลง แต่ถ้าเปลี่ยนค่า σ จะทำให้รูปร่างของโค้ง

เปลี่ยนแปลง





อย่างไรก็ตาม การแจกแจงปกติมีคุณลักษณะที่สำคัญอีกดังนี้

(1) $f(x) > 0$ สำหรับทุกค่าของ x

(2) $f(x)$ จะลดลงเรื่อย ๆ (ถึง 0) ถ้า x มีค่าห่างจาก μ เพิ่มขึ้น ๆ (ถึง∞) และ $f(x)$ จะมีค่ามากที่สุดเป็น $\frac{1}{\sqrt{2\pi}}$ ถ้า $x = \mu$

(3) $f(x)$ สมมาตรที่ μ นั่นคือ $f(\mu + x) = f(\mu - x)$ สำหรับทุกค่า x

(4) เมื่อ $x = \mu$ แล้ว $f(x)$ จะสูงสุด ดังนั้น μ ซึ่งเป็นค่าคาดหวังนั้นจะเป็นฐานนิยม (Mode) และมัธยมฐาน (Median) ของการแจกแจงปกติด้วย

(5) ส่วนโค้งจะแคบเข้า ถ้า σ ลดลง นั่นคือ ถ้ากำหนดให้ $k > 0$ แล้วพื้นที่ใต้โค้งซึ่งอยู่ระหว่าง $\mu - k$ และ $\mu + k$ จะเพิ่มขึ้น ถ้า σ ลดลง

(6) พื้นที่ใต้โค้งระหว่าง $\mu - \sigma$ กับ $\mu + \sigma$ เท่ากับ .6826; $\mu - 2\sigma$ กับ $\mu + 2\sigma$ เท่ากับ .9545 และ $\mu - 3\sigma$ กับ $\mu + 3\sigma$ เท่ากับ .9973

แต่ที่เราใช้กันบ่อย ๆ คือ พื้นที่ใต้โค้งระหว่าง $\mu - 1.96\sigma$ กับ $\mu + 1.96\sigma$ เท่ากับ .95 และ $\mu - 2.58\sigma$ กับ $\mu + 2.58\sigma$ เท่ากับ .99

(7) จุดกลับ (Point of inflection) ของโค้งอยู่จุด $(\mu-\sigma, \frac{1}{\sigma\sqrt{2\pi e}})$ และ $(\mu+\sigma, \frac{1}{\sigma\sqrt{2\pi e}})$

การแจกแจงปกติได้พัฒนามาดังนี้ ในปี 1733 Abraham de Moivre (1667 - 1745) เป็นคนแรกที่นำเอาการแจกแจงปกติมาใช้ประมาณความน่าจะเป็นของการแจกแจงทิวนามที่มีค่า p ต่าง ๆ และ n มีค่าใหญ่ และยังใช้ประมาณความน่าจะเป็นของการแจกแจงปัวซองได้ด้วย ต่อมาปี 1809 Karl F Gauss (1777 – 1855) พบว่าหากมีเลขอยู่หลาย ๆ จำนวนซึ่งได้มาจากการวัด เช่น ความสูงของคนหลาย ๆ คน จำนวนเหล่านี้เมื่อนำมาจัดเข้าในตารางความถี่และแสดงการแจกแจงแล้วกราฟของการแจกแจงมักจะอยู่ในรูปของสมการโค้งปกติเสมอ ในระยะแรก ๆ การแจกแจงแบบนี้มักเรียกกันว่าการแจกแจงแบบเกาส์ (Gaussian Distribution) การแจกแจงนี้เองที่ Gauss

ได้นำมาใช้เป็นรากฐานของทฤษฎีว่าด้วยความเคลื่อนคลาด (Theory of Error) ของการวัดหรือการนับจำนวนใด ๆ

ทฤษฎีที่สำคัญที่ทำให้การแจกแจงปกติมีความสำคัญตามที่กล่าวมาแล้วในตอนแรก คือ ทฤษฎีว่าด้วยขีดจำกัดเข้าส่วนกลาง (Central Limit Theorem, CLT) ซึ่งกล่าวไว้คร่าว ๆ ว่า "ฟังก์ชั่นแจกแจงน่าจะเป็นใด ๆ สามารถประมาณได้ด้วยฟังก์ชั่นแบบปกติ เสมอ ในเมื่อมีเงื่อนไขบางประการในขีดจำกัด"

การประมาณความน่าจะเป็นจากการแจกแจงทวินามและปัวซองนั้นเป็นกรณีพิเศษของ CLT การนำไปใช้ที่สำคัญที่สุดของการแจกแจงปกติและ CLT ก็คือในการสุ่มตัวอย่างซึ่ง CLT ยืนยันว่า "การแจกแจงของค่าเฉลี่ยจากตัวอย่างจะมีลักษณะเป็นโค้งปกติเสมอ หากขนาดของตัวอย่างใหญ่พอ"

ตัวอย่าง ถ้า X เป็นเส้นผ่าศูนย์กลางภายในของกระบอกสูบเครื่องยนต์ ซึ่งมีการแจกแจงปกติที่มีค่าคาดหวัง μ และความแปรปรวน 1 ถ้า X ไม่ได้มาตรฐานที่ตั้งไว้ ผู้ผลิตต้องสูญเสีย (loss) และถ้าสมมติว่า P เป็นกำไรต่อหนึ่งกระบอกสูบเป็นฟังก์ชั่นของ X ดังนี้

$$
\begin{aligned}
P &= C_1, \quad 10 \leq x \leq 12 \\
&= -C_2, \quad x < 10 \\
&= -C_3, \quad x > 12
\end{aligned}
$$

ในเมื่อ C_1, C_2 และ C_3 เป็นค่าคงที่ซึ่งเป็นเงินที่ได้กำไร

ดังนั้นกำไรคาดหวัง (ต่อกระบอกสูบ) เขียนได้เป็นดังนี้

$$
\begin{aligned}
E(P) &= C_1 \int_{10}^{12} f(x)\,dx - C_2 \int_{-\infty}^{x} f(x)\,dx - C_3 \int_{12}^{\infty} f(x)\,dx \\
&= C_1 \{P(x < 12) - P(x < 10)\} - C_2 \{P(X < 10)\} \\
&= C_1 \{P(z < 12 - \mu) - P(z < 10 - \mu)\} \\
&\quad - C_2 \{P(z < 10 - \mu)\} - C_3 \{P(z > 12 - \mu)\} \\
&= (C_1 + C_3) P(z < 12 - \mu) - (C_1 + C_2) P(z < 10 - \mu) - C_3
\end{aligned}
$$

ในกระบวนการผลิตสามารถปรับค่า μ ได้ ถ้าจะให้ได้กำไรคาดหวังสูงสุดต้องกำหนดค่า μ เหมาะ ๆ ซึ่งจะกำหนดได้โดยการแก้สมการ

$$\frac{dE(P)}{d\mu} = 0$$

นั่นคือ $0 = (C_1 + C_3)\{-P(z < 12 - \mu)\} - (C_1 + C_2)\{-P(z < 10 - \mu)\}$

$$\mu = 11 - \frac{1}{2} \ln \left(\frac{C_1 + C_3}{C_1 + C_2}\right)$$

การแจกแจงปกติมาตรฐาน (Standard Normal Distribution) **เมื่อเราต้องการจะคำนวณ** ความน่าจะเป็นที่ตัวแปรเชิงสุ่ม x จะมีค่าอยู่ในช่วง [a, b] เราหาได้ดังนี้

$$P(a \le X \le b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{x-\mu}{\sigma})^2} dx$$

สำหรับ μ และ σ ที่กำหนดให้

เราจะเห็นได้ว่ายากที่จะคำนวณ แต่เราก็มีวิธีง่าย ๆ **ที่จะคำนวณความน่าจะเป็นที่ต้อง**การด้วยการใช้ตารางที่เตรียมไว้สำหรับฟังก์ชั่นพิเศษแบบหนึ่ง **นั่นคือฟังก์ชั่นหนาแน่นปกติมาตร**ฐานและด้วยตัวแปลงง่าย ๆ เราสามารถแสดงว่าการแจกแจง**ปกติก็เป็นเช่นเดียวกับฟังก์ชั่นหนา**แน่นปกติมาตรฐานดังต่อไปนี้

ถ้าให้ **X** เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงปกติ $N(x; \mu, \sigma^2)$ ที่มี $\mu = 0$ และ $\sigma^2 = 1$ แล้ว การแจกแจงปกตินี้จะได้ชื่อว่าการแจกแจงปกติมาตรฐาน ซึ่งจะมีฟังก์ชั่น ดังนี้

$$f(z) = 1/\sqrt{2\pi}\ e^{-z^2/2} \quad ; -\infty < z < \infty$$

ถ้าตัวแปรเชิงสุ่ม x มีการแจกแจงปกติ $N(x; \mu, \sigma^2)$ แล้วตัวแปรเชิงสุ่มใหม่ Z **ที่กำหนด**ไว้ว่า $Z = (X-\mu)/\sigma$ นั้นจะมีการแจกแจงปกติมาตรฐาน $N(x; 0, 1)$

จากฟังก์ชั่นหนาแน่นปกติมาตรฐาน เราจะเห็นว่าไม่มีพารามิเตอร์อยู่เลย **และจากฟังก์ชั่นนี้ได้สร้าง**ตารางที่คำนวณไว้ซึ่งเรียกว่า ตารางปกติ เราจะใช้ตารางนี้คำนวณความน่า**จะเป็นปกติ**ได้ แต่ต้องแปลงการแจกแจงปกติให้เป็นการแจกแจงปกติมาตรฐานเสียก่อน

ตัวอย่าง จำนวนผลิตภัณฑ์นมที่ขายได้ในวันหนึ่ง ๆ มีการแจกแจงปกติที่มี $\mu = 200$ บาท และ $\sigma = 30$ บาท

จงหาโอกาสที่จะขายได้มากกว่า 225 บาท และระหว่าง 150 กับ 200 บาท

ให้ X เป็นจำนวนผลิตภัณฑ์นมที่ขายได้ในวันหนึ่ง ๆ ซึ่งแปลงให้เป็น**มาตรฐานได้ด้วย**ตัวแปลง $Z = (X-u)/\sigma$

ดังนั้น $P(X > 225) = P(\frac{X-200}{30} > \frac{225-200}{30}$

$= P(Z > 0.83) = 1 - P(Z < 0.83)$

$= 1 - .7967 = .2033$

$P(150 < X < 2200) = P(\frac{150-200}{30} < Z < \frac{200-200}{300})$

$= P(-1.67 < Z < 0)$

$= P(0 < Z < 1.67) = 0.4525$

จงหาจำนวนขาย k ที่ทำให้ $P(X < k) = 0.25$

$$P(X < k) = P\left(Z < \frac{k - 200}{30}\right) = 0.25$$

จากตาราง $P(Z < -0.67) = .2514$

ดังนั้น $\frac{k - 200}{30} = -.67$

$k = 200 - .67\,(30) = 179.90$

(ก) การประมาณการแจกแจงทวินามด้วยการแจกแจงปกติ ได้กล่าวมาแล้วว่า การแจกแจงปกติเป็นขีดจำกัด (limiting form) ของการแจกแจงหลายอย่าง ในกรณีที่เป็นขีดจำกัดของการแจกแจงทวินามที่มี n ขนาดใหญ่ (โดยปกติ $n \geq 20$)แสดงได้ดังนี้

$$\sum_{x=s_1}^{s_2} \binom{n}{x} p^x q^{n-x} \cong \int_{a_1}^{a_2} \frac{1}{\sqrt{2\pi}} e^{-1/2x^2} dx$$

ในเมื่อ $a_1 = (s_1 - nP)/\sqrt{nPQ}$ และ $a_2 = (s_2 - nP)/\sqrt{NPQ}$

การประมาณนี้จะถูกต้อง ถ้า $n \to \infty$ และ $P \to \frac{1}{2}$ ถึงแม้ว่า n เล็ก และ P ไม่ใกล้ 0 หรือ 1 เกินไป การประมาณก็ยังใช้ได้ แต่จะให้ได้ผลเป็นที่น่าพอใจแล้ว nP และ nQ ควรจะมีค่ามากกว่า 5

G.W. Cochran ได้แนะนำค่าของ n และ p ในการประมาณความน่าจะเป็นของการแจกแจงทวินามด้วยการแจกแจงปกติไว้ในหนังสือเทคนิคการสุ่มตัวอย่าง (Sampling Techniques, (1953) ดังนี้

ค่าของ P	ขนาดของ n อย่างน้อย
.5	30
.4 หรือ .6	50
3 หรือ .7	80
.2 หรือ .8	200
.1 หรือ .9	600
05 หรือ .95	1400

ในการประมาณการแจกแจงทวินามด้วยการแจกแจงปกตินั้น เป็นการประมาณการแจกแจงของตัวแปรเชิงสุ่มไม่ต่อเนื่องด้วยตัวแปรเชิงสุ่มต่อเนื่อง โดยที่ $P(X = a) = 0$ ถ้าตัวแปรเชิงสุ่ม X ต่อเนื่อง แต่ $P(X = a)$ อาจจะมากกว่า 0 ถ้า X ไม่ต่อเนื่อง ดังนั้น จำเป็นต้องแก้ไขจุดปลาย

(end points) ด้วย ± 0.5 การแก้ไขเช่นนั้นเรียกว่า การแก้ไขของการต่อเนื่อง (Correction for continuity) นั่นคือ a_1 และ a_2 จะเป็น

$$a_1 = (s_1 - nP - .5)/\sqrt{nPQ} \text{ และ } a_2 = (s_2 - nP + .5)/\sqrt{nPQ}$$

ตัวอย่างเช่น (1) $P(X = a) \cong P(a-nP-.5)/\sqrt{nPQ} < Z < (a-nP+.5)/\sqrt{nPQ}$

(2) $P(a \leq X \leq b) \cong P[(a-nP-.5)/\sqrt{nPQ} < Z < (b-nP+.5)/\sqrt{NPQ}]$

ตัวอย่าง วิธีการผลิตอย่างหนึ่งทำให้เกิดของเสีย 5% ถ้าสุ่มของที่ผลิตออกมา 100 ชิ้น จงหาความน่าจะเป็นที่มีของเสียอย่างมาก 4 ชิ้น

$\mu = np = 100(.05) = 5$

$\sigma^2 = npq = 100(.05)(.95) = 4.75$

ดังนั้น $P(X \leq 4) = P(0 \leq x \leq 4) = P[(0 - .5 - 5)/\sqrt{4.75} < Z < (4 + .5 - 5)/\sqrt{4.75}]$

$= P(-2.52 < Z < -0.23) = .4070$

แต่ความน่าจะเป็นจริง ๆ จากการแจกแจงทวินามจะเป็น .4360

(ข) การประมาณการแจกแจงปัวซองด้วยการแจกแจงปกติ การแจกแจงปัวซองใช้ประมาณ การแจกแจงทวินามในเมื่อ n โต และ P น้อย ๆ และการแจกแจงปกติใช้ประมาณการแจกแจงทวินามเมื่อ n โต (np ไม่น้อยเกินไป) ซึ่งจะกลายเป็นว่าการแจกแจงปกติประมาณปัวซองเมื่อ α (ค่าเฉลี่ย) โตพอ Laplace ได้พิสูจน์ไว้ว่า

$$\lim_{\alpha \to \infty} \frac{e^{-\alpha}\alpha^x}{x!} \cong 1/\sqrt{2\pi\alpha}\, e^{\left(\frac{x-\mu}{\sqrt{\alpha}}\right)^2}$$

ดังนั้นเราจะได้ว่า $\sum_{x=n_1}^{n_2} \frac{e^{-\alpha}\alpha^x}{x!} \cong \int_{a1}^{a2} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx$

ในเมื่อ $a_1 = (n_1 - \alpha - .5)/\sqrt{\alpha}$ และ $a_2 = (n_2 - \alpha + .5)/\sqrt{\alpha}$

การประมาณจะเที่ยงตรง ถ้า α มีค่าใหญ่พอ

ตัวอย่าง จำนวนของที่ขายได้ในวันหนึ่ง ๆ มีการแจกแจงปัวซองที่มีค่าเฉลี่ย 64 ชิ้นต่อวัน จงหาจำนวนของที่ทำให้โอกาสที่ขายได้มากกว่าจำนวนนั้นเป็น 0.8

โดยที่ $\alpha = 64$ ชิ้น, k = จำนวนของที่ต้องการ และทราบว่า $P(X > k) = .80$

ดังนั้น $P(X > k) = P(X \geq k + 1) = P(Z \geq (k + 1 - .5 - 64)/\sqrt{64}) = 0.80$

แต่จากตารางปกติเราทราบว่า $P(Z > -0.84) = 0.80$

ดังนั้น $-0.84 = \frac{k+1-64-.5}{\sqrt{64}}$

$$k = 64-8(0.84)-5 = 58$$

(ค) การประมาณการแจกแจงปาสคาลด้วยการแจกแจงปกติ เมื่อ k ซึ่งเป็นจำนวนครั้งที่ประสบความสำเร็จนั้นมีขนาดโต เราสามารถประมาณการแจกแจงปาสคาลด้วยการแจกแจงปกติ โดยอาศัยทฤษฎีขีดจำกัดสู่ส่วนกลางได้ดังนี้

$$\sum_{y=n_1}^{n_2} \binom{y+k-1}{k-1} P^k(1-P)^y \cong \int_{a_1}^{a_2} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

ในเมื่อ

$$a_1 = (n_1 - kQ/P - .5)/\sqrt{kQ/P^2} \quad \text{และ} \quad a_2 = (n_2 - kQ/P + .5)/\sqrt{kQ/P^2}$$

(3) การแจกแจงเอกซ์โพเนนท์ (Exponential distribution) ตัวแปรเชิงสุ่ม x จะเป็นตัวแปรเชิงสุ่ม เอกซ์โพเนนท์ ถ้ามีฟังก์หนาแน่นน่าจะเป็นดังนี้

$$f(x) = \alpha e^{-\alpha x};\ x > 0$$

ตัวแปรเชิงสุ่มปัวซองกับเอกซ์โพเนนท์มีความสัมพันธ์กัน กล่าวคือ ถ้า y เป็นตัวแปรเชิงสุ่มปัวซองที่เป็นจำนวนครั้งของการเกิดขึ้นของเหตุการณ์หนึ่งในคาบเวลาที่กำหนด แล้ว x ซึ่งเป็นตัวแปรเชิงสุ่มที่เป็นคาบเวลาของการรอคอยของการเกิดขึ้นครั้งแรก หรือสองครั้งติดต่อกันของเหตุการณ์จะมีการแจกแจงเอกซ์โพเน้นท์ดังนี้

$$f(x) = \alpha e^{-\alpha x};\ x > 0$$

การแจกแจงเอกซ์โพเน้นท์มีบทบาทสำคัญในทฤษฎีความเชื่อมั่น (Reliability Theory) และปัญหาการรอคอย (Waiting Time or Queueing Problems) การแจกแจงนี้มีคุณสมบัติที่สำคัญ

(1) ฟังก์ชันแจกแจงกำหนดไว้ดังนี้

$$F(x) = 1-e^{-\alpha x};\ x > 0$$

(2) ค่าคาดหวังและความแปรปรวนเป็น

$$E(X) = 1/\alpha,\ V(X) = 1/\alpha^2$$

ตัวอย่าง ให้ y เป็นตัวแปรเชิงสุ่มซึ่งเป็นจำนวนที่ต่อโทรศัพท์ไม่ติด และมีการแจกแจงปัวซอง ถ้าโดยเฉลี่ยต่อไม่ติด 2 ครั้ง ในหนึ่งชั่วโมง และให้ x เป็นเวลาระหว่างการต่อไม่ติด 2 ครั้งติดต่อกันแล้วเราจะได้ว่า

$$f(y) = e^{-2}\, 2^y/y!\quad y = 0, 1, 2, \ldots\ldots\ldots$$
$$f(x) = 2e^{-2x};\quad x > 0$$

ถ้าสมมติว่าโทรศัพท์ต่อไม่ติด ความน่าจะเป็นที่โทรศัพท์ต่อไม่ติดจะเกิดขึ้นอีกภายใน 2 ชั่วโมงหน้านี้จะคำนวณได้เป็น

$$P(0 \leq x \leq 2) = \int_0^2 2e^{-2x}\,dx = 1 - e^{-4}$$
$$= 0.982$$

(4) การแจกแจงแกมมะ (Gamma Distribution) การแจกแจงแกมมะเป็นรูปทั่วไปของการแจกแจงเอกซโพเนนท์ นั่นคือ ถ้าเราทราบว่า x เป็นตัวแปรเชิงสุ่มที่แทนเวลาที่คอยเหตุการณ์จะเกิดขึ้นครั้งแรก แล้ว x มีการแจกแจงเอกซโพเนนท์ สมมติว่า เรากำหนดตัวแปรเชิงสุ่ม T ซึ่งเป็นเวลาทั้งหมดที่รอคอยเหตุการณ์จะเกิดขึ้น k ครั้ง แล้วสามารถแสดงได้ว่า

$$T = X_1 + X_2 + \ldots\ldots + X_k$$

ในเมื่อ X_i เป็นเวลาที่รอคอยของการเกิดขึ้นครั้งที่ i (i = 1, 2,...., k) โดยที่ X_i เป็นอิสระต่อกันและ

$$f(t) = \frac{\alpha}{\Gamma(k)} (\alpha t)^{k-1} e^{-\alpha t}\quad t > 0$$

นั่นคือเวลาทั้งหมด T ของการเกิดขึ้น k ครั้งติดต่อกันจะมีฟังก์ชั่นหนาแน่นน่าจะเป็น $f(t)$ โดยกำหนดว่าเวลาที่รอคอยภายในการเกิดขึ้น 2 ครั้งติดต่อกันเป็นแบบเอกซ์โพเนนท์ ฟังก์ชั่นหนาแน่น $f(t)$ เรียกว่า ฟังก์ชั่นหนาแน่นน่าจะเป็นแบบแกมมะ

ตัวแปรเชิงสุ่ม X ที่มีการแจกแจงแบบแกมมะจะมีฟังก์ชั่นหนาแน่นดังนี้

$$f(x) = \frac{\alpha}{\Gamma(k)} (\alpha x)^{k-1} e^{-\alpha x};\ x > 0,\ \alpha > 0,\ k > 1$$

ในเมื่อ $\Gamma(k)$ เรียกแกมมะฟังก์ชั่น (Gamma Function) ซึ่งกำหนดไว้ดังนี้

$$\Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x}\,dx,\ k > 0$$
$$= (k-1)\,\Gamma(k-1) = (k-1)!$$

สำหรับ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

การแจกแจงแกมมะมีคุณสมบัติที่น่าสนใจดังนี้

(1) ถ้า k = 1 แล้ว $f(x) = \alpha e^{-\alpha x}$ จะเป็นฟังก์ชั่นหนาแน่นเอกซโพเนนท์

(2) ถ้า k เป็นจำนวนเต็มบวก แล้วฟังก์ชั่นแจกแจงแกมมะสามารถเขียนได้ในรูปฟังก์ชั่นแจกแจงปัวซอง

$$f(x) = 1 - \sum_{s=0}^{k-1} e^{-\alpha x} (\alpha x)^s / s!;\ x > 0$$

ความสัมพันธ์นี้พูดได้อีกอย่างว่า การแจกแจงปัวซองจะเกี่ยวกับจำนวนการเกิดขึ้นของเหตุการณ์

บางอย่างระหว่างคาบเวลาที่กำหนด แต่การแจกแจงแกมมะจะเป็นการแจกแจงของเวลาที่ต้องการให้ได้จำนวนการเกิดขึ้นของเหตุการณ์ตามที่กำหนดไว้ เช่น ถ้า x เป็นจำนวนการเกิดขึ้นของเหตุการณ์ A ระหว่าง (O,t) ภายใต้สภาวการณ์ที่เหมาะ X จะมีการแจกปัวซองที่มี $E(X) = \alpha t$ เมื่อ α เป็นจำนวนคาดหวังของการเกิดขึ้นของ A ระหว่างช่วงเวลาหนึ่งหน่วย ถ้า T เป็นเวลาที่ต้องการใช้สังเกตการเกิดขึ้นของ A จำนวน k ครั้ง เราจะได้ความสัมพันธ์ดังกล่าว นั่นคือ

$$F(t) = P(T \leq t) = 1 - P(T > t)$$

$$= 1 - P\{\text{การเกิดขึ้นของเหตุการณ์ A น้อยกว่า k ครั้งใน } (0, t)\}$$

$$= 1 - P(X < k)$$

$$= 1 - \sum_{s=0}^{k-1} e^{-\alpha t}(\alpha t)^{s}/s!$$

(3) ถ้า $\alpha = \frac{1}{2}$, $k = \nu/2$, และ ν เป็นค่าเต็มบวก แล้วตัวแปรเชิงสุ่ม x จะมีการแจกแจงไคสแควร์ (Chi - Square) ที่องศาแห่งความเป็นอิสระ (degrees of freedom) เท่ากับ ν นั่นคือ

$$f(x) = \frac{\frac{1}{2}}{\Gamma(\nu/2)} (x/2)^{\nu/2-1} e^{-x/2}, \quad x > 0$$

(4) ค่าคาดหวังและความแปรปรวนของการแจกแจงแกมมะหาได้ดังนี้

$E(X) = k/\alpha$, $V(X) = k/\alpha^2$

การแจกแจงแกมมะ บางครั้งเรียกว่า Erlangian or Pearson's Type III Distribution

การแจกจงแกมมะมีส่วนสัมพันธ์กับการแจกแจงไคสแคว์ที่มีบทบาทสำคัญในทฤษฎีสถิติและนอกจากนั้นยังมีการประยุกต์ต่าง ๆ ในปัญหาการตัดสินใจอีกด้วย

สรุปการเปรียบเทียบระหว่างการแจกแจงต่าง ๆ

1. กระทำการทดลองแบบ เบอร์นุยยี (Bernoulli Trials) ที่เป็นอิสระต่อกัน

(1) ตัวแปรเชิงสุ่ม : จำนวนครั้งของการเกิดขึ้นของเหตุการณ์ที่สนใจ (A) ในจำนวนของการทดลองที่กำหนด (n)

การแจกแจง : ทวินาม

(2) ตัวแปรเชิงสุ่ม : จำนวนการทดลองแบบ เบอร์นุยยี (Bernoulli Trials) ที่จะต้องทำทั้งหมดเพื่อให้การเกิดขึ้นครั้งแรกของเหตุการณ์

การแจกแจง : เรขาคณิต

(3) ตัวแปรเชิงสุ่ม : จำนวนการทดลองแบบ เบอร์นุยยี (Bernoulli Trials) ที่จะทำทั้งหมดเพื่อให้การเกิดขึ้นของเหตุการณ์ A เป็นครั้งที่ k

การแจกแจง : ปาสคาลหรือทวินามนิเสธ

2. กลุ่ม (ไม่จำกัด) ของตัวแปรเชิงสุ่มที่เป็นตัวแปรเชิงสุ่มปัวซอง

(4) ตัวแปรเชิงสุ่ม : จำนวนของการเกิดขึ้นของเหตุการณ์ A ในช่วงเวลาที่กำหนด
การแจกแจง : ปัวซอง

(5) ตัวแปรเชิงสุ่ม : เวลาที่ต้องรอยคอยจนกว่าเหตุการณ์ A เกิดขึ้นครั้งแรก
การแจกแจง : เอกซโพเน้นท์

(6) ตัวแปรเชิงสุ่ม : เวลาที่ต้องรอคอยจนกว่าเหตุการณ์ A เกิดขึ้นครั้งที่ k
การแจกแจง : แกมมะ

หมายเหตุ ลองเปรียบเทียบความคล้ายคลึงกันระหว่าง (1) กับ (3), (2) กับ (4) และ (3) กับ (6)

(5) **การแจกแจงบีทะ** (Beta distribution) ตัวแปรเชิงสุ่มบางตัวเช่น สัดส่วนของเสียอาจจะมีอยู่ในช่วง (0, 1) เท่านั้น และความน่าจะเป็นของช่วงย่อย ๆ (Subinterval) ที่เท่ากันของช่วง (0, 1) นั้นจะไม่เท่ากัน (ถ้าเท่ากันแล้วตัวแปรเชิงสุ่มนั้นจะมีการแจกแจงยูนิฟอร์ม) ตัวแปรเชิงสุ่มที่มีลักษณะเช่นนี้จะมีการแจกแจงบีทะซึ่งมีฟังก์ชั่นหนาแน่นน่าจะเป็นดังนี้

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}; \quad 0 < x < 1; \; \alpha, \beta > 0$$

$$= \frac{1}{B(\alpha,\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}$$

ในเมื่อ $B(\alpha,\beta)$ เรียกว่า บีทะฟังก์ชั่น (Beta function) กำหนดไว้ดังนี้

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

คุณสมบัติที่น่าสนใจของการแจกแจงบีทะ ก็คือ

(1) $E(X) = \alpha/(\alpha+\beta)$

$V(X) = \alpha\beta/(\alpha+\beta)^2(\alpha+\beta+1)$

(2) ฐานนิยมของการแจกแจงเท่ากับ $(\alpha-1)/(\alpha+\beta-2)$

(3) รูปร่างการแจกแจงนี้ขึ้นอยู่กับค่าของ α และ β สำหรับค่าของ α และ β ต่าง ๆ กันก็จะให้กราฟต่าง ๆ กันด้วย การแจกแจงจะสมมาตรก็ต่อเมื่อ $E(X) = .5$

(4) ถ้า $\alpha = 1$ และ $\beta = 1$ แล้วการแจกแจงนี้จะกลายเป็นการแจกแจงยูนิฟอร์ม

เมื่อไม่นานมานี้ การแจกแจงบีทะใช้ประยุกต์ในทฤษฎีตัดสินใจโดยเฉพาะในการกำหนดความน่าจะเป็นก่อนการทดลองและหลังทดลองให้แก่สภาวะการณ์นอกบังคับ (States of nature)

(6) **การแจกแจงไวบุล** (Weibull Distribution) การแจกแจงนี้ใช้ประยุกต์ในทฤษฎีความเชื่อถือได้บ่อยที่สุด ถ้าให้ X เป็นตัวแปรเชิงสุ่มแทนอายุของการบริการ (Life length or time to failure) ของส่วนประกอบในระบบ และมีฟังก์ชั่นหนาแน่นน่าจะเป็นต่อไปนี้แล้วเราจะเรียกว่า X เป็นตัวแปรเชิงสุ่มไวบุล

$$f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}}; \; x > 0; \; \alpha, \beta > 0$$

ถ้าให้ Z เป็นอัตราการเสีย (Instantaneous failure rate) ซึ่งกำหนดไว้ดังนี้

$$Z(x) = f(x)/R(x) = f(x)/(1-F(x))$$

ในเมื่อ R(x) เป็นความเชื่อถือได้ของส่วนประกอบที่กำหนด ณ เวลา x นั่นคือ

$$R(x) = P(X > x) = 1 - P(X \leqslant x)$$
$$= 1 - F(x) = e^{-\alpha x^{\beta}}$$

ดังนั้น ความเชื่อถือได้ของส่วนประกอบที่กำหนดจึงเป็นความน่าจะเป็นที่ส่วนประกอบจะทำงานปกติอย่างน้อยเท่ากับเวลาที่กำหนด ภายใต้เงื่อนไขของการทดลองที่กำหนดไว้)

ฉะนั้น $Z(x) = \alpha\beta x^{\beta-1};\ \alpha > 0, \beta > 0,\ x > 0$

คุณสมบัติที่น่าสนใจของการแจกแจงไวบุล มีดังนี้

(1) ค่าคาดหวังและความแปรปรวนของ X เป็นดังนี้

$$E(X) = \alpha^{-1/\beta}\Gamma(1 + 1/\beta)$$
$$V(X) = \alpha^{-2/\beta}\{\Gamma(1 + 2/\beta) - [\Gamma(1 + 1/\beta)]^2\}$$

(2) เมื่อ $\beta = 1$ การแจกแจงนี้จะเป็นการแจกแจงเอกซโพเนนท์ และเมื่อ $\beta > 1$ กราฟของการแจกแจงจะคล้ายรูประฆัง

(3) เมื่อ $\beta = 2$ เราจะได้การแจกแจงเรย์เล่ซ์ (Rayleigh distribution) ซึ่งใช้ประยุกต์ทางทฤษฎีเสียงและการทดสอบอาวุธ (Noise Theory and Weapon Testing) นั่นคือฟังก์ชั่นแบบเรย์เล่ซ์กำหนดไว้ดังนี้

$$f(x) = 2\alpha x e^{-\alpha x^2} \qquad x > 0, \alpha > 0$$

ซึ่งมีค่าคาดหวังและความแปรปรวนเป็น

$$E(X) = (1/2\alpha)\sqrt{\pi}$$
$$V(X) = \frac{1}{\alpha}(1 - \pi/4)$$

(7) **การแจกแจงพาเรโท** (Pareto Distribution) การแจกแจงนี้ตั้งตามชื่อของ Vilfredo Pareto นักสังคมวิทยาและเศรษฐศาสตร์ชาวอิตาลีซึ่งความสนใจในการแจกแจงรายได้ (Income Distribution) การแจกแจงนี้ใช้อธิบายปรากฏการต่าง ๆ เช่น การแจกแจงระดับรายได้, ขนาดของเมือง, และความถี่ของคำ เป็นต้น การแจกแจงพาเรโท กำหนดไว้ดังนี้

$$f(x) = (\alpha/x_0)(x_0/x)^{\alpha+1};\ x > x_0;\ x_0,\ \alpha > 0$$

ค่าเฉลี่ย มัธยฐาน และความแปรปรวน จะเป็นดังนี้

$$E(X) = \frac{\alpha x_0}{\alpha - 1} \qquad \alpha > 1$$
$$X_{med} = X_0 \sqrt[\alpha]{2}$$
$$V(X) = \alpha x_0^2 / (\alpha - 1)^2(\alpha - 2),\ \alpha > 2$$

(8) **การแจกแจงไคสแควร์** (Chi-Square X^2 Distribution) ตัวแปรเชิงสุ่ม x จะมีการแจกแจงไคสแควร์ ถ้าฟังก์ชันหนาแน่นน่าจะเป็นของมันกำหนดไว้ดังนี้

$f(x) = [1/\Gamma(k/2)](x/2)^{k/2-1} e^{-x/2}$, $x > 0$, $k \geq 1$ ในเมื่อ k เป็นค่าคงที่เป็นจำนวนเต็ม และเรียกว่า องศาความเป็นอิสระของ X

ค่าคาดหวัง และความแปรปรวน จะเป็น

$E(X) = k$, $V(X) = 2k$

การแจกแจงไคสแควร์นี้มีประโยชน์อย่างมากในสถิติอนุมาน และตารางของการแจกแจงนี้ได้ใช้กันแพร่หลาย ค่าในตารางนั้นเป็นค่าของ X ที่ทำให้ $P(X > x) = \alpha$ โดยที่ k เป็นองศาความเป็นอิสระที่ระบุไว้ในแถวตั้งทางซ้ายมือ และ α เป็นความน่าจะเป็นที่ระบุไว้ในแถวนอนบน

ตัวแปรเชิงสุ่มไคสแควร์กับตัวแปรเชิงสุ่มปกติมาตรฐานมีความสัมพันธ์กันดังนี้ (1) ถ้า $Z_1, Z_2, \ldots, Z_k$ เป็นตัวแปรเชิงสุ่มปกติมาตรฐานที่เป็นอิสระ แล้ว $X^2 = Z_1^1 + Z_2^2 + \ldots + Z_k^2$ จะเป็นตัวแปรเชิงสุ่มที่มีการแจกแจงไคสแควร์ ด้วยองศาความเป็นอิสระ k;(2) ถ้า Z มีการแจกแจงปกติมาตรฐาน และ x มีการแจกแจงไคสแควร์ที่มีองศาความเป็นอิสระ 1 แล้ว $P(X > a^2) = P(Z < -a) + P(Z > a) = 2P(Z > a)$ และ (3) ถ้า y มีการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ k และ k โตพอแล้ว $Z = \sqrt{2y} - \sqrt{2k-1}$ จะมีการแจกแจงปกติมาตรฐาน

นอกจากนี้การแจกแจงไคสแควร์ยังมีความสัมพันธ์กับการแจกแจงปัวซองดังนี้ ให้ x เป็นตัวแปรเชิงสุ่มปัวซองที่มีพารามิเตอร์ λ และให้ r เป็นเลขจำนวนเต็มที่ทำให้

$P(X \geq r) \leq \alpha < P(X \geq r-1)$

แล้วจะได้ว่า

$P(X \geq r) = P(y \leq 2\lambda)$

ในเมื่อ y เป็นตัวแปรเชิงสุ่มไคสแควร์ที่มีองศาความเป็นอิสระ 2r

นั่นคือถ้าเราต้องการหาค่า r ที่ทำให้ 2λ อยู่ระหว่างจุด 100 α % ล่าง (Lower 100 α Percent Points) ของการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ 2r และ 2(r - 1) แล้ว r จะเป็นจุด 100 α % บน (Upper 100 α percent points) ของการแจกแจงปัวซองที่มีค่าเฉลี่ย λ

ในทำนองเดียวกัน ถ้า r เป็นค่าที่ทำให้

$P(X \leq r) \leq \alpha < P(X \leq r+1)$

แล้วเราจะได้ว่า ค่า

$P(X \leq r) = P(Z \geq 2\lambda)$

ในเมื่อ Z มีการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ 2(r + 1)

เมื่อเราต้องการหาค่า r ที่ทำให้ 2λ อยู่ระหว่างจุด 100 % บนของการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ 2r และ 2(r + 1) แล้ว r จะเป็นจุด 100 α % ล่างของการแจกแจงปัวซอง

ที่มีค่าเฉลี่ย λ

ตัวอย่าง (1) จงหาจุด 5% บนของการแจกแจงปัวซองที่มีค่าเฉลี่ย 0.9

เราจะหา r ที่ทำให้จุด 5% ล่างของการแจกแจงไคสแควร์ (ที่มีองศาความเป็นอิสระ 2r) หรือ $\chi^2(2r)$ นั้นมากกว่า 2λ = 1.8 และจุด 5% ล่างของการแจกแจง $\chi^2(2(r-1))$ น้อยกว่า 1.8 จากตารางไคสแควร์เราพบว่า r = 4

(1) จงหาจุด 10% ล่างของการแจกแจงปัวซองที่มีค่าเฉลี่ย 5.0

เราจะหา r ที่ทำให้จุด 10% บนของการแจกแจง χ^2 (2r) น้อยกว่า 2 = 10.0 และจุด 10% บนของการแจกแจง $\chi^2(2(r+1))$ มากกว่า 10.0 จากตารางไคสแควร์ เราพบว่า r = 2

การแจกแจงไคสแควร์นี้ F.R. Helmert เป็นผู้พัฒนา และ K. Pearson เป็นผู้นำไปประยุกต์เกี่ยวกับการทดสอบสมมติฐาน

(9) **การแจกแจงแบบที** (Student t distribution) ตัวแปรเชิงสุ่ม x จะมีการแจกแจงแบบทีถ้าฟังก์ชั่นกำหนดไว้ดังนี้

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\sqrt{k\pi}} \cdot \left(1+\frac{x^2}{k}\right)^{-(k+1)/2}, \quad -\infty < x < \infty$$

การแจกแจงแบบทีนี้มีความสัมพันธ์กับการแจกแจงไคสแควร์และปกติมาตรฐานดังนี้ "ถ้า Z มีการแจกแจงปกติมาตรฐาน U มีการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ k และถ้า z และ U เป็นอิสระกัน แล้วตัวแปรเชิงสุ่ม $X = Z/\sqrt{U/k}$ จะมีการแจกแจงแบบทีด้วยองศาความเป็นอิสระ k"

William Sealy Gosset เป็นผู้พัฒนาการแจกแจงแบบทีในปี 1908 ซึ่งเขาเขียนผลการพัฒนาชื่อ The Probable Error of a Mean ในวารสาร Biometrika โดยใช้นามปากกา "Student"

กราฟของการแจกแจงแบบทีมีลักษณะเป็นรูประฆัง คล้ายโค้งปกติมาตรฐาน แต่แบนกว่าเล็กน้อย ถ้า k ยิ่งโตกราฟก็จะใกล้โค้งมาตรฐานยิ่งขึ้น นั่นคือเมื่อ $k \to \infty$ การแจกแจงแบบทีจะเป็นแบบปกติมาตรฐาน

$$\lim_{k\to\infty} f(x_i) \to \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

ค่าคาดหวังและความแปรปรวนกำหนดไว้ดังนี้

$E(X) = 0$, $V(X) = \frac{k}{k-2}$, $k > 2$

ซึ่งเราจะเห็นว่า $V(X) > 1$

การแจกแจงนี้มีบทบาทต่อสถิติอนุมานและการวิจัยเป็นอย่างมากนั้นเมื่อใช้อ้างอิงเกี่ยวกับค่าเฉลี่ยประชากรแบบที่ใช้กันอย่างกว้างขวาง

(10) **การแจกแจงแบบเอฟ** (Snedecor F - Distribution) ถ้า U และ V เป็นตัวแปรเชิงสุ่มที่เป็นอิสระกันและมีการแจกแจงไคสแควร์ด้วยองศาแห่งความเป็นอิสระ m และ n ตามลำดับ

แล้ว $X = \dfrac{U/m}{V/n}$ จะเป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบเอฟที่มีองศาแห่งความเป็นอิสระ m และ n ดังนี้

$$f(x) = \frac{\Gamma(m/2+n/2)}{\Gamma(m/2)\Gamma(n/2)}(m/n)^{m/2}\left(1+\frac{m}{n}x\right)^{-(m+n)/2}x^{m/2-1};\ x>0$$

G.W. Snedecor ได้เป็นผู้พัฒนาการแจกแจงแบบเอฟขึ้นมาและตั้งชื่อให้เป็นเกียรติแก่ Ronald Fisher ว่าการแจกแจงเอฟ เพราะ Fisher ได้เป็นผู้ศึกษาครั้งแรกในปี 1924 การแจกแจงแบบเอฟมีบทบาทในสถิติการทดลองมาก และมีคุณสมบัติที่น่าสนใจดังนี้

(ก) ค่าคาดหวังและความแปรปรวนหาได้ดังนี้

$E(X) = n/(n-2);\ n > 2$

$V(X) = \dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)},\ n > 4$

(ข) การแจกแจงแบบเอฟจะเบ้ทางขวา ความเบ้จะลดลงถ้าจำนวนองศาความเป็นอิสระเพิ่มขึ้น

(ค) การแจกแจงแบบเอฟมีคุณสมบัติของส่วนกลับดังนี้ "ตัวแปรเชิงสุ่ม $y = 1/x$ จะมีการแจกแจงแบบเอฟด้วยองศาความเป็นอิสระ n และ m ซึ่งกลับกันกับการแจกแจงของ x ที่มีองศาความเป็นอิสระ m และ n นั่นคือ

$P(F^{(m,n)} < X) = P(F^{(n,m)} > 1/X)$

หรือ $P(F^{(m,n)} > X) = P(F^{(n,m)} < 1/X)$"

(ง) การแจกแจงแบบที กับแบบเอฟมีส่วนสัมพันธ์กันดังนี้ "ถ้า X มีการแจกแจงแบบทีด้วยองศาความเป็นอิสระ k แล้ว X^2 จะมีการแจกแจงแบบเอฟด้วยองศาความเป็นอิสระ (1,k) นั่นคือ

$P(F^{(1,n)} > a^2) = P(t^{(n)} > a) + P(t^{(n)} < -a) = 2P(t^{(n)} > a)$"

(จ) การแจกแจงไคสแควร์กับแบบเอฟมีความสัมพันธ์กันดังนี้ "ถ้า y มีการแจกแจงแบบเอฟด้วยองศาความเป็นอิสระ m, ∞ แล้ว mX จะมีการแจกแจงไคสแควร์ด้วยองศาความเป็นอิสระ m นั่นคือ

$P(F^{(m,\infty)} > a/m) = P(X^{2(m)} > a)$"

(ฉ) การแจกแจงทวินามกับแบบเอฟมความสัมพันธ์กันดังนี้ "ให้ X เป็นตัวแปรเชิงสุ่มทวินามที่มีพารามิเตอร์ n และ P และให้ r เป็นค่าเลขเดิมบวกที่ทำให้

$P(X \geq r) \leq \alpha < P(X \geq r-1)$ แล้วเราจะได้ว่า

$P(X \geq r) = P\left(Y \geq \dfrac{r}{n+1-r}\,\dfrac{1-P}{P}\right)$

ในเมื่อ Y เป็นตัวแปรเชิงสุ่มแบบเอฟด้วยองศาความเป็นอิสระ 2(n + 1-r) และ 2r

สำหรับ r เมื่อให้เป็นค่าที่ทำให้ $P(X \leq r) \leq \alpha < P(X \leq r+1)$ แล้วเราจะได้อีกว่า

$$P(X \leq r) = P\left(Z \geq \frac{n-r}{r+1} \cdot \frac{P}{1-P}\right)$$

ในเมื่อ Z เป็นตัวแปรเชิงสุ่มแบบเอฟด้วยองศาความเป็นอิสระ $2(r+1)$ และ $2(n-r)$

จากความสัมพันธ์นี้เราสามารถใช้ประมาณค่าและทดสอบสมมติฐานเกี่ยวกับค่า P นี้ได้

(11) **การแจกแจงปกติแบบล๊อก** (Lognormal Distribution) ถ้าให้ x เป็นตัวแปรเชิงสุ่มที่เป็นบวก $(x > 0)$ และให้ $Y = \ln X$ โดยที่ Y มีการแจกแจงปกติ แล้ว X จะมีการแจกแจงปกติแบบล๊อก ดังนี้

$$f(x) = (1/X\sigma\sqrt{2\pi})\, e^{-(1/2\sigma^2)(\ln x - \mu)^2}; \quad x>0;\ \mu>0;\ \sigma>0$$

ค่าคาดหวังและความแปรปรวน จะเป็นดังนี้

$$E(X) = e^{\mu + \sigma^2/2}$$

$$V(X) = e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2}$$

$$E(\ln X) = \mu,\ V(\ln X) = \sigma^2$$

การแจกแจงแบบนี้จะมีฐานนิยมเดียว แต่เบ้ทางขวา เมื่อ σ^2 มีค่าน้อย แล้วรูปร่างของการแจกแจงจะใกล้กับโค้งปกติ

(12) การแจกแจงปกติสองมิติ (Bivariate Normal Distribution) ตัวแปรเชิงสุ่ม (X, Y) ซึ่งเป็นแบบต่อเนื่องจะมีการแจกแจงปกติสองมิติ ถ้าฟังก์ชั่นหนาแน่นน่าจะเป็นร่วมกำหนดไว้ดังนี้

$$f(x,y) = (1/2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})\ \exp\left[-\frac{1}{2(1-\rho)^2}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\cdot\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}\right]$$

การแจกแจงนี้มีค่าเฉลี่ย ความแปรปรวน และสัมประสิทธิ์สหสัมพันธ์ ดังนี้

$$E(X) = \mu_n,\ E(y) = \mu$$

$$E(y/x) = \mu_y + \rho(\sigma_y/\sigma_x)(X-\mu_x)$$

$$V(X) = \sigma_x^2,\ V(y) = \sigma_y^2$$

$$\rho_{xy} = \rho,\quad Cov(x, y) = \rho\sigma_x\sigma_y$$

เราจะเห็นได้ว่า E(y/x) เป็นฟังก์ชั่นเชิงเส้น (Linearfunction) ของ X