บทที่ 2. ความน่าจะเป็นเบื้องต้น

Probability is the very Guide of Life.

Bishop Joseph Butler

ความน่าจะเป็นมีบทบาทสำคัญในกระบวนการตัดสินใจ (Decision-making Process) ไม่ว่า ปัญหานั้นจะเป็นปัญหาที่ประสบในทางธุรกิจ เศรษฐกิจ ดาราศาสตร์ สังคมศาสตร์ หรือในชีวิตประจำวันของตัวเราเอง การตัดสินใจในสภาวะที่มีข้อมูลข่าวสารพร้อมมูล (Perfect information) หรือมีข้อเท็จจริงที่จำเป็นพร้อมมูลนั้นเห็นจะเกิดขึ้นไม่บ่อยนัก ส่วนมากเรามักจะทำการตัดสินใจในสภาวะที่มีข้อมูลข่าวสารไม่พร้อม หรือในสภาวะที่ไม่แน่นอน (Uncertainty) ดังนั้นความน่าจะเป็นจึงเข้าไปมีบทบาทในกระบวนการตัดสินใจในฐานะตัวแทนของความแน่นอน (Certainty) หรือตัวแทนของความรู้ข้อเท็จจริงที่สมบูรณ์ (Complete Knowledge) นั่นเอง

ในสถิติอนุมาน (Statistical Inference) ความน่าจะเป็นเข้าไปมีบทบาทอย่างสำคัญ เพราะเป็นรากฐานของทฤษฎีและการประยุกต์ทางสถิติ งานที่สำคัญของนักสถิติก็คือทำการสรุปผลหรือทำการอ้างอิงจากการทดลองหรือปรากฏการณ์ซึ่งมีความไม่แน่นอนเกี่ยวข้องอยู่ จากทฤษฎีความน่าจะเป็นจะทำให้นักสถิติสามารถวางหลัก (Generalize) จากสิ่งที่รู้ไปหาสิ่งที่ไม่รู้ และการวางหลักก็สามารถทำด้วยความเชื่อมั่นสูง ๆ ได้ ดังนั้น ทฤษฎีความน่าจะเป็นจึงเป็นเครื่องมือที่สำคัญ ในการวิเคราะห์ทางสถิติ

2.1 ความหมายของความน่าจะเป็น

ความน่าจะเป็น (Probability) ก็เป็นอีกชนิดหนึ่งที่เราใช้กันบ่อย ๆ โดยเฉพาะในวงการพนัน แต่เราก็ยังใช้กันอยู่จำกัดมาก ทั้ง ๆ ที่เป็นคำที่มีบทบาทในชีวิตประจำวันของเราเป็นอย่างมาก และคำที่มีความหมายเช่นเดียวกับคำนี้ก็คือ โอกาสหรือโคลก (chance) และความเป็นไปได้ (Possibility) ความน่าจะเป็นมีความหมายอยู่สองประการคือ

ประการแรก ความน่าจะเป็นในฐานะที่เป็นศาสตร์ (Science) ซึ่งเป็นวิชาที่ศึกษาเกี่ยวกับการทดลองเชิงสุ่ม (Random Experiment) หรือการทดลองที่ไม่สามารถบอกผลการทดลองได้ด้วยความแน่ใจ แต่จะบอกได้ด้วยความไม่แน่ใจหรือด้วยความน่าจะเป็น เช่นการทดลองโดยการโยนเหรียญ ทอดลูกเต๋า, นับจำ นวนเครื่องบินขึ้นลงในสนามบินดอนเมือง ในช่วงระยะเวลาหนึ่ง เป็นต้น

ประการที่สอง ความน่าจะเป็นในฐานะที่เป็นตัวเลข (Numbers) ซึ่งเป็นมาตรงวัดเชิงปริมาณ (Quantitative Measure) ของความไม่แน่นอน ที่เหตุการณ์จากการทดลองเชิงสุ่มจะเกิดขึ้น ถ้าเหตุการณ์ที่สนใจจะต้องเกิดขึ้นแน่นอน แล้วความน่าจะเป็นที่เหตุการณ์จะเกิดขึ้นจะเป็น 1.00 แต่ถ้าการเกิดขึ้นของเหตุการณ์ที่สนใจไม่มีทางเป็นไปได้ แล้วความน่าจะเป็นที่เหตุการณ์

นั้นจะเกิดขึ้นจะเป็น 0.00 ดังนั้นถ้าเหตุการณ์ต่าง ๆ ที่การเกิดขึ้นของมันเป็นไปได้ แต่ไม่แน่ใจแล้วความน่าจะเป็นที่จะเกิดขึ้นของเหตุการณ์จะอยู่ระหว่าง 0.00 และ 1.00 เหตุการณ์ที่มีโอกาสเกิดขึ้นมาก ก็จะมีความน่าจะเป็นใกล้ 1.00 แต่ถ้ามีโอกาสเกิดขึ้นน้อยก็จะมีความน่าจะเป็นห่างจาก 1.00 หรืออยู่ใกล้ 0.00 นั่นเอง เราจะเห็นได้ว่าความน่าจะเป็นมีสเกลในเทอมของสัดส่วนเป็น (0,1) หรือในเทอมของเปอร์เซนต์ (ซึ่งใช้กันบ่อย ๆ) จะเป็น (0,100) เช่น ความน่าจะเป็นของการเกิดหน้าคู่จากการทอดลูกเต๋าหนึ่งลูกจะเป็น 0.5 หรือ 50% ความน่าจะเป็นที่จะสอบผ่านวิชาสถิติจะเท่ากับ 0.9 หรือ 90% เป็นต้น

2.2 การทดลอง (Experiment)

จากนิยามของความน่าจะเป็นซึ่งเกี่ยวข้องกับการทดลอง ดังนั้นเราจำเป็นต้องนิยามการทดลองกันก่อนดังนี้การทดลองเป็นกระบวนการที่กระทำขึ้นภายใต้กลุ่มของเงื่อนไขอย่างหนึ่ง และกระบวนการนั้นสามารถทำซ้ำ ได้ภายใต้เงื่อนไขเดียวกัน เมื่อสิ้นสุดลงกระบวนการก็จะให้ผลออกมาซึ่งจะเรียกว่าผลทดลอง (Outcomes) และผลทดลองที่เป็นไปได้จะมีตั้งแต่หนึ่งอย่างขึ้นไป การทดลองจะแบ่งออกเป็น 2 ชนิด ดังนี้

ก. การทดลองแบบกำหนดได้ (Deterministic Experiment) เป็นการทดลองที่เราทราบด้วยความแน่ใจว่าผลทดลองที่จะเกิดขึ้นจากการทดลองแต่ละครั้งเป็นอะไร การทดลองแบบนี้เป็นการทดลองที่เราคุ้นเคยกันมาก เช่นโยนของขึ้นไป เราจะบอกได้ล่วงหน้าว่าของชิ้นนั้นจะตกลงพื้นอย่างแน่นอน ถ้ากู้เงินจากธนาคาร 10,000 บาท ภายใน 1 ปี เราก็ทราบอย่างแน่นอนว่า เราต้องเสียดอกเบี้ย 1,500 บาท (ถ้าอัตราดอกเบี้ยในปีนั้นไม่เปลี่ยนแปลง)

ข. การทดลองเชิงสุ่ม (Random or Nondeterministic Experiment) การทดลองแบบนี้เป็นการทดลองที่เป็นจริงหรือสมมุติขึ้นมา และต้องมีคุณสมบัติ 3 ประการ ต่อไปนี้

(1) ก่อนที่จะทำการทดลองนั้น ผลทดลองที่จะเกิดขึ้นจะทราบด้วยความไม่แน่ใจนั่นคือผลทดลองโคลก (Chance factors) เข้าไปเกี่ยวข้องด้วย

(2) กลุ่มผลทดลอง (Sample Space) หรือเซทของผลทดลองที่เป็นไปได้ทั้งหมดจะทราบด้วยความแน่ใจ ก่อนที่จะทำการทดลอง และ

(3) ในสภาวการณ์ (Condition) เดียวกันนั้น การทดลองสามารถจะทำซ้ำ ๆ ได้

ข้อมูลดิบที่เกี่ยวกับการวิเคราะห์ทางสถิติ ส่วนมากก็ได้มาจากการทดลองแบบนี้

ตัวอย่าง **การทดลอง** : ทอดลูกเต๋า (Fair die) และสังเกตหน้าที่เกิดขึ้นเราจะเห็นได้ว่า (1) ก่อนการทอดลูกเต๋าที่จะเกิดขึ้นได้ทั้งหมดนั้น เป็น $S = \{1, 2, 3, 4, 5, 6\}$ และ (3) การทอดลูกเต๋านี้สามารถจะทำซ้ำ ๆ ได้ในสภาวการณ์เดิม

ดังนั้นการทอดลูกเต๋านี้ จึงเป็นการทดลองเชิงสุ่ม

ตัวอย่างของการทดลองเชิงสุ่มอื่น ๆ

(1) โยนสตางค์อันหนึ่ง 2 ครั้ง และสังเกตจำนวนหัวที่เกิดขึ้น

(2) โยนสตางค์อันหนึ่งจนกว่าจะเกิดหัว 2 ครั้ง และสังเกตจำนวนครั้งที่จะโยน

(3) สังเกตอายุของหลอดทีวีที่ติดตั้งในมหาวิทยาลัยรามคำแหง

(4) สังเกตจำนวนของที่ไม่ได้มาตรฐานที่ผลิตออกมาในคาบเวลา 24 ชั่วโมง จากกระบวนการผลิตแบบอัตโนมัติ

2.3 กลุ่มผลทดลอง และเหตุการณ์ (Sample Space and Events)

(ก) กลุ่มผลทดลอง (Sample Space) เป็นเซทของผลทดลอง (Sample points, points, outcomes) ที่เป็นไปได้ทั้งหมดจากการทำการทดลองเชิงสุ่ม เราจะใช้ S หรือ Ω, หรือ $\mathcal{U}$ แทนกลุ่มผลทดลอง

สำหรับกลุ่มผลทดลองมีชื่อเรียกทางภาษาอังกฤษอย่างอื่นอีกคือ Outcome space, Event space หรือ Universal space และกลุ่มผลทดลองนั้น แบ่งออกได้เป็น 2 ประเภท คือ

(1) กลุ่มผลทดลองไม่ต่อเนื่อง (Discrete Sample Space) เป็นกลุ่มผลทดลองที่ประกอบด้วยผลทดลองที่เป็นไปได้ทั้งหมดมีจำนวนจำกัด หรือไม่จำกัดแต่แจงนับได้ (Finite or Countable infinite)

(2) กลุ่มผลทดลองต่อเนื่อง (Continuous Sample Space) เป็นกลุ่มผลทดลองที่มีผลทดลองที่เป็นไปได้ทั้งหมดมีจำนวนไม่จำกัด และแจงนับไม่ได้ (Uncountably infinite)

ตัวอย่าง จากตัวอย่างของการทดลองเชิงสุ่มอื่น ๆ ที่ผ่านมา เราจะได้กลุ่มผลทดลอง S ดังนี้

(1) $S = \{TT, TH, HT, HH\} = \{0, 1, 2\}$ S เป็นกลุ่มผลทดลองไม่ต่อเนื่อง

(2) $S = \{2, 3, 4, \ldots\}$ S เป็นกลุ่มผลทดลองไม่ต่อเนื่อง

(3) $S = \{T / T \geqslant 0\}$, T เป็นอายุการใช้งานของหลอดทีวีซึ่งมีหน่วยเป็นชั่วโมง S เป็นกลุ่มผลทดลองต่อเนื่อง

(4) $S = \{0, 1, 2, \ldots\}$, D เป็นจำนวนของที่ไม่ได้มาตรฐานที่ผลิตออกมาในคาบเวลา 24 ชั่วโมง S เป็นกลุ่มผลทดลองไม่ต่อเนื่อง

(ข) เหตุการณ์ (Event) เป็นเซทย่อย (Subsets) ของกลุ่มผลทดลองหรือเป็นเซทของผลทดลองที่เป็นไปได้จากที่การทำการทดลองเชิงสุ่ม

เหตุการณ์ที่น่าสนใจมีดังนี้

(1) เกตุการณ์แบบง่าย (Simple or Elementary Event) เป็นเหตุการณ์หรือเซทที่ประกอบด้วยผลทดลองที่เป็นไปได้เพียงหนึ่งเท่านั้น

(2) เหตุการณ์ประกอบ (Compound Event) เป็นเหตุการณ์ที่ประกอบด้วยผลทดลองที่เป็นไปได้มากกว่าหนึ่ง

(3) เหตุการณ์ที่เกิดแน่นอน (Certain or Sure Event) เป็นเหตุการณ์ที่ประกอบด้วย

ผลทดลองที่เป็นไปได้ทั้งหมด เหตุการณ์นี้ก็คือกลุ่มผลทดลองนั่นเอง

(4) เหตุการณ์ที่เกิดไม่ได้ (Impossible, Null or Void Event) เป็นเหตุการณ์ที่ไม่มีผลทดลองที่เป็นไปได้เลย

ตัวอย่าง **การทดลอง** : โยนสตางค์ 3 ครั้ง และสังเกตผลทอลอง (คือหน้า) ที่จะเกิดขึ้นจากการโยน

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

ถ้าเรากำหนด E_1, E_2, E_3 และ E_4 เป็นเหตุการณ์ที่มีคุณสมบัติดังนี้

E_1 : เหตุการณ์ที่ได้หัวทั้งสามครั้ง

E_2 : เหตุการณ์ที่ได้หัวอย่างน้อย 2 ครั้ง

E_3 : เหตุการณ์ที่ได้หัว 3 ครั้ง หรือน้อยกว่า

E_4 : เหตุการณ์ที่ได้หัวมากกว่า 3 ครั้ง

แล้วเราจะได้ E_1 = { HHH } เป็นเหตุการณ์แบบง่าย

E_2 = { HHH, HHT, HTH, THH } เป็นเหตุการณ์ประกอบ

E_3 = { HHH, HHT, HTH, THH, HTT, THT, TTH, TTT } = S เป็นเหตุการณ์ที่เกิดแน่นอน

E_4 = { } = ϕ เป็นเหตุการณ์ที่เกิดขึ้นไม่ได้

2.4 การรวมตัวของเหตุการณ์ (Event Operations)

เนื่องจากเหตุการณ์เป็นเซท การรวมของเหตุการณ์จึงเหมือนกับการรวมตัวของเซทที่น่าสนใจมีดังนี้

(1) **ผลรวมของเหตุการณ์** (Union or Sum of Events) ถ้า A และ B เป็นเหตุการณ์ แล้วผลรวมของ A และ B หรือ AUB จะเป็นเหตุการณ์ที่ประกอบด้วยผลทดลองที่เป็นไปได้ที่อยู่ในเหตุการณ์ A หรือ B นั่นคือ

$$A \cup B = \{ s_i / s_i \in A \text{ หรือ } s_i \in B \}$$

ในเมื่อ s_i เป็นผลทดลองที่เป็นไปได้ ส่วน<u>หรือ</u> (or) ในที่นี้มีความหมายเป็น " และ/หรือ (and/or) "

(2) ผลร่วมของเหตุการณ์ (Intersection or Product of Events, or Joint events): A และ B เป็นเหตุการณ์ แล้วผลร่วมของ A และ B หรือ A∩B จะเป็นเหตุการณ์ที่เกิดขึ้นก็ต่อเมื่อทั้ง A และ B เกิดขึ้นพร้อม ๆ กัน นั่นคือ

$$A \cap B = \{ s_i / s_i \in A \quad \& \quad s_i \in B \}$$

สองเหตุการณ์ A และ B จะเรียกว่า **ไม่มีผลร่วมกัน** (Mutually exclusive or Disjoint event) ถ้าทั้งสองเหตุการณ์เกิดขึ้นพร้อมกัน นั่นคือ $A \cap B = \phi$

(3) **ส่วนเติมเต็มของเหตุการณ์** (Complement or Absolute Complement of Events) ถ้า A เป็นเหตุการณ์ แล้วส่วนเติมเต็มของ A, $\bar{A}$, จะเป็นเหตุการณ์ที่จะเกิดขึ้นก็ต่อเมื่อเหตุการณ์ A ไม่ได้เกิดขึ้น นั่นคือ

$$\bar{A} = \{s_i | s_i \in S \ , \quad s_i \notin A\}$$

(4) **ผลต่างของสองเหตุการณ์** (Difference of Two Events) ถ้า A และ B เป็นเหตุการณ์ แล้วผลต่างของสองเหตุการณ์ A และ B หรือ A-B จะเป็นเหตุการณ์ที่ประกอบด้วยผลทดลองที่เป็นไปได้ที่อยู่ในเหตุการณ์ A แต่ไม่อยู่ในเหตุการณ์ B นั่นคือ

$$A - B = \{s_i | s_i \in A \ , \quad s_i \notin B\} = A \cap \bar{B}$$

(5) **ส่วนแบ่งของเหตุการณ์** (Partition of Event) เหตุการณ์ $A_1, A_2, \ldots, A_n$ จะประกอบเป็นส่วนแบ่งของเหตุการณ์ B ถ้าสอดคล้องกับคุณสมบัติต่อไปนี้

(1) $A_i \subseteq B$ (i = 1, 2, ….. , n) นั่นคือ A_i เป็นเหตุการณ์ที่มีผลทดลองทั้ง หมดอยู่ใน B

(2) $A_i \cap A_j = \phi$ (i, j = 1, 2, …… , n, $i \neq j$) นั่นคือ A_i และ A_j ไม่มีผลร่วมกัน

(3) $A_1 \cup A_2 \cup \ldots \cup A_n = B$ นั่นคือผลรวมของเหตุการณ์ $A_1, A_2, \ldots A_n$ จะเป็นเหตุการณ์ B นั่นเอง

การรวมตัวของเหตุการณ์นั้นสามารถเป็นแผนภาพที่เรียกว่า Venn Diagram ซึ่งตั้งชื่อตาม John Venn (1834-1923) นักตรรกวิทยาชาวอังกฤษ แต่ถ้าแผนภาพได้แสดงผลทดลองที่เป็นไปได้ของกลุ่มผลทดลองด้วยจุดภายในวงกลม เราจะเรียกแผนภาพนั้นว่า Euler Diagram ซึ่งตั้งชื่อตาม Leonhard Euler(1707-1783) นักคณิตศาสตร์และฟิสิกส์ชาวสวิส

ตัวอย่าง **การทดลอง** : โยนลูกเต๋าครั้งหนึ่ง และสังเกตหน้าที่ขึ้น

ดังนั้น S = {1, 2, 3, 4, 5, 6,}

ถ้ากำหนดเหตุการณ์ต่าง ๆ ดังนี้

A : เกิดแต้มคู่ = {2, 4, 6}

B : ได้แต้มไม่เกิน 3 = {1, 2, 3}

C : เกิดแต้มคี่ = {1, 3, 5}

D : ได้แต้มเท่ากับ 6 หรือน้อยกว่า = {1, 2, 3, 4, 5, 6}

E : ได้แต้มน้อยกว่า 1 = { }

แล้วเราจะได้การรวมตัวของเหตุการณ์ดังต่อไปนี้

ผลรวม	$A \cup B$	=	{1, 2, 3, 4, 6}	$A \cup C$	=	{1, 2, 3, 4, 5, 6,}	
	$B \cup C$	=	{1, 2, 3, 5}	$D \cup E$	=	{1, 2, 3, 4, 5, 6,}	= S

ผลร่วม : $A \cap B = \{2\}$ $A \cap C = \{\ \}$

$B \cap C = \{1, 3\}$ $D \cap E = \{\ \}$

ส่วนเติมเต็ม : $\bar{A} = \{1, 3, 5\} = C$ $\bar{B} = \{4, 5, 6\}$

$\bar{D} = \{\ \} = E$

ผลต่าง : $A - B = \{4, 6\}$, $B - C = \{2\}$

ส่วนแบ่ง : พิจารณาเหตุการณ์ A และ C เราจะเห็นว่า (1) $A \cap C = \{\ \}$ และ (2) $A \cup C = S$ ดังนั้น A และ C เป็นส่วนแบ่งของ S

ตัวอย่าง **กำหนดให้** F เป็นเหตุการณ์ของครอบครัวทั้งหมดในไทย

(ก) ให้ A_1 เป็นครอบครัวที่มีรายได้น้อยกว่า 10,000 บาท ; A_2 มีรายได้ระหว่าง 10,000 บาท และ 50,000 บาท ; A_3 มีรายได้ระหว่าง 50,000 บาท และ 100,000 บาท A_4 มีรายได้ระหว่าง 100,000 บาท และ 200,000 บาท ; และ A_5 มีรายได้อย่างน้อย 200,000 บาท ต่อปี

ถ้าเราให้ $S = \{A_1, A_2, A_3, A_4, A_5\}$ เราจะเห็นได้ว่า

(1) $A_i \subset F;\ i = 1, 2, 3, 4, 5,$

(2) $A_i \cap A_j = \phi,\ i \neq j$

(3) $A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 = F$

จึงสรุปได้ว่า S เป็นส่วนแบ่งของ F

(ข) ถ้าให้ B_1 เป็นเหตุการณ์ของครอบครัวที่ผู้มีรายได้หลักมีอายุน้อยกว่า 20 ปี ; B_2 มีอายุระหว่าง 20 ปี ถึง 30 ปี ; B_3 ระหว่าง 30 ถึง 40 ; B_4 ระหว่าง 40 ถึง 50 ; B_5 ระหว่าง 50 ถึง 65 ; และ B_6 อายุตั้งแต่ 65 ปีขึ้นไป

เราจะเห็นได้ว่า

(1) $B_i \subset F;\ i = 1, 2, 3, 4, 5, 6,$

(2) $B_i \cap B_j = \phi,\ i \neq j$

(3) $B_1 \cup B_2 \cup B_3 \cup B_4 \cup B_5 \cup B_6 = F$

เราจึงสามารถสรุปได้ว่า $B_1, B_2 \ldots\ldots\ldots, B_6$ ประกอบเป็นส่วนแบ่งของ F

(ค) ถ้าเราพิจารณาผลร่วมของเหตุการณ์ B_i และ A_k หรือเหตุการณ์ $B_i \cap A_k$; $i = 1, 2, 3, 4, 5, 6,$ และ $k = 1, 2, 3, 4, 5,$ เราจะเห็นได้ว่า

(1) $(B_i \cap A_k) \subset F$

(2) $(B_i \cap A_k) \cap (B_j \cap A_l) = \phi;\ i \neq j,\ k \neq l$

(3) $(B_1 \cap A_1) \cup (B_2 \cap A_1) \cup \ldots\ldots\ldots \cup (B_6 \cap A_5)$

$$= \left(\bigcup_{i=1}^{6} B_i\right) \cap \left(\bigcup_{k=1}^{5} A_k\right)$$

$$= \bigcup_{i=1}^{6} \bigcup_{k=1}^{5} (B_i \cap A_k) = F$$

ดังนั้นเหตุการณ์ $B \cap A_k$ เหล่านี้จะประกอบเป็นส่วนแบ่งของ F แต่ในกรณีนี้เราจะเรียกว่า ส่วนแบ่งร่วม (Cross Partition)

ตัวอย่าง เครื่องอีเล็กโทรนิคได้รับการทดสอบเพื่อสังเกตอายุของการใช้งาน สมมุติว่าเป็น t ดังนั้นกลุ่มผลทดลองจะเป็นดังนี้ $S = \{t/t \geq 0\}$

ถ้าให้ A, B และ C เป็นเหตุการณ์ที่กำหนดไว้ดังนี้.-

$A = \{t/t < 100\}$; $B = \{t/50 \leq t \leq 200\}$; $C = \{t/t > 150\}$

แล้วจะได้ว่า $A \cup B = \{t/t \leq 200\}$, $B \cup C = \{t/t \geq 50\}$

$A \cap B = \{t/50 \leq t < 100\}$

$B \cap C = \{t/150 < t \leq 200\}$

$A \cup C = \{t/t < 100 \text{ or } t > 150\}$ $\qquad A \cap C = \emptyset$

$\bar{A} = \{t/t \geq 100\}$; $\bar{C} = \{t/t \leq 150\}$

$A - B = \{t/t < 50\}$; $B - C = \{t/50 \leq t \leq 150\}$

สำหรับเหตุการณ์ A, B และ C นั้น ไม่เป็นส่วนแบ่งของ S เพราะ $A \cap B \neq \emptyset$ หรือ $B \cap C \neq \emptyset$

โดยที่ (1) A, B และ C มีผลรวมเป็น S นั่นคือ $A \cup B \cup C = S$; (2) A, B, C ต่างก็เป็นเซทย่อยของ S นั่นคือ $A \subset S$, $B \subset S$, หรือ $C \subset S$ เราจะเรียกเหตุการณ์ A, B และ C ว่าเป็นเหตุการณ์ผลรวม (Exhaustive events)

2.5 วิธีนับจำนวนผลทดลองในเหตุการณ์หรือกลุ่มผลทดลอง (Methods of Enumeration)

ในการทดลองซึ่งจะให้ผลทดลองออกมานั้น ถ้าเราต้องการทราบจำนวนผลทดลองที่เป็นไปได้ทั้งหมด เราก็ทำการแจงนับเอา แต่การทดลองบางอย่างจะทำการแจงนับยาก ต้องอาศัยหลักคำนวณ และก็เช่นเดียวกันสำหรับเหตุการณ์ที่เราสนใจซึ่งเราต้องการทราบจำนวนผลทดลอง เรามีกฎหรือหลักดังนี้

(1) **หลักการคูณ** (Multiplication Principle) ถ้ามีวิธีการ (Procedures) k วิธีการ และวิธีการที่ 1, 2,....., k สามารถกระทำได้ $n_1, n_2, \ldots, n_k$ วิธีตามลำดับแล้ววิธีการที่จะประกอบด้วยวิธีการที่ 1, ตามด้วยวิธีการที่ 2, ... , ตามด้วยวิธีการที่ k จะสามารถกระทำได้ $n_1 (n_2) (n_3) \ldots (n_k)$ วิธี

ตัวอย่าง หีบ 365 ใบ จะมีวิธีที่จะใส่ลูกบอล 30 ลูก ลงในหีบ โดยหีบใดหีบหนึ่งมีไม่เกิน 1 ลูก ได้กี่วิธี?

บอลลูกแรกจะเลือกใส่ได้ 365 วิธี (หีบ)

บอลลูกที่สองจะเลือกได้ 364 วิธี

บอลลูกที่ 30 จะเลือกได้ [365 - (30 - 1)] วิธี

ดังนั้น วิธีที่จะใส่บอล 30 ลูก ลงในหีบ 365 ใบ จะมีวิธีทำได้เท่ากับ

$365 (364) (363) \ldots\ldots [365 - (30 - 1)]$ วิธี

(2) **หลักการบวก** (Addition Principle) ถ้ามีวิธีการอยู่ k วิธี และวิธีการที่ 1, 2, 3, ..., k สามารถกระทำได้ $n_1, n_2, \ldots\ldots, n_k$ ตามลำดับ แล้วจำนวนวิธีที่สามารถกระทำวิธีการ 1 หรือ วิธีการ 2 หรือ ...... หรือวิธีการใด ๆ สามารถกระทำร่วมกันได้

ตัวอย่าง ถ้าเราวางแผนที่จะไปเที่ยวเชียงใหม่ และกำลังตัดสินใจระหว่างการเดินทางโดยรถไฟและรถยนต์ เรารู้ว่าทางรถไปทางเชียงใหม่มีทางเดียว แต่ทางรถยนต์มี 3 ทาง

ดังนั้น วิธีที่จะไปเที่ยวเชียงใหม่จะมี 1 + 3 = 4 วิธี

(3) **การจัดเรียงอันดับ** (Permutation) การจัดเรียงอันดับเป็นการจัดเรียงสิ่งของทั้งหมดหรือเพียงบางส่วน โดยคำนึงถึงอันดับที่

ตัวอย่าง วิธีที่จะจัดเรียงอักษร ก, ข, และ ค มีดังนี้

(1) จัดเรียงทั้งหมด

(2) จัดเรียงเพียง 2 ตัว

(1)	ก	ข - ค	กขค	(2) ก	ข	กข
		ค - ข	กคข		ค	กค
	ข	ก - ค	ขกค	ข	ก	ขก
		ค - ก	ขคก		ค	ขค
	ค	ก - ข	คกข	ค	ก	คก
		ข - ก	คขก		ข	คข

ในการจัดเรียงอันดับ เรามีกฎสำหรับคำนวณวิธีการดังนี้

กฎที่ 1 สิ่งของที่ต่างกัน n สิ่ง ถ้านำมาจัดเรียงกันทีละ r สิ่ง โดยคำนึงถึงอันดับ **แต่ไม่มีการแทนที่** แล้ว วิธีการจัดเรียงทั้งหมดจะเป็น

$$P(n, r) = \frac{n!}{(n-r)!} \quad ; r \leq n$$

หมายเหตุ (1) ถ้า n เป็นเลขโดดที่เป็นบวก (Positive Integer) เรานิยาม n! ซึ่งเรียกว่า n - factorial ได้ดังนี้

$$n! = n(n-1)(n-2) \ldots (2)(1) \; ; 0! = 1$$

ตัวอย่างเช่น $5! = 5(4)(3)(2)(1) = 120$

(2) $P(n,r)$ อาจเขียนได้เป็น ${}^{n}P_{r}$ หรือ nPr หรือ(n)r เช่น $P(5,2) = {}^{5}P_{2} = 5P2 = (5)2$

ตัวอย่าง เลข 0 ถึง 9 ถ้าจะนำมาประกอบเป็นจำนวนเลข 2 หลัก จะได้กี่วิธี (ไม่ให้เลขซ้ำกัน)

เลข 0 ถึง 9 จะมีตัวเลขอยู่ 10 ตัว

ดังนั้น $P(10, 2) = 10!/(10-2) = 90$ วิธี

กฎที่ 2 จำนวนการจัดเรียงอันดับของสิ่งของที่ต่างกัน n สิ่ง ซึ่งนำมาจัดเรียงทีละ r สิ่ง โดยการแทนที่ จะได้

$$P_1(n, r) = n^r$$

ตัวอย่าง จำนวนเลข 2 หลัก ที่ประกอบขึ้นจากเลข 0 ถึง 9 จะได้กี่จำนวน?

$$P_1(10, 2) = 10^2 = 100$$

กฎที่ 3 การจัดเรียงอันดับแบบวงกลม (Circular Permutations) : จำนวนวิธีของการจัดเรียงอันดับของสิ่งของที่ต่างกัน n สิ่ง โดยจัดเรียงเป็นวงกลม จะได้

$$P_2(n, n) = (n-1)!$$

ตัวอย่าง นักศึกษา 3 คน นั่งเรียงกันเป็นวงกลม จะได้กี่วิธี?

$$P_2(3, 3) = (3-1)! = 2$$

กฎที่ 4 ถ้าสิ่งของ n สิ่ง แบ่งเป็นออกเป็นพวกที่เหมือนกัน k พวก, พวกแรกมี n_1 สิ่ง, พวกที่สองมี n_2 สิ่ง, ... , พวกที่ k มี n_k สิ่ง แล้วจำนวนการจัดเรียงอันดับกำหนดไว้ดังนี้

$$P_3(n, n) = n!/\{(n_1!)(n_2!)\ldots\ldots(n_k!)\} \; ; n_1 + n_2 \ldots\ldots + n_k = n$$

ตัวอย่าง **วิธีที่จะจัดเรียงอันดับของคำว่า** "Statistics" ได้กี่วิธี?

Statistics มี S จำนวน 3 ตัว, t 3 ตัว, i 2 ตัว, และ a, c อย่างละตัว

ดังนั้น จำนวนจัดเรียงอันดับจะเป็น

$$P_3(10, 10) = \frac{10!}{(3!)(3!)(2!)(1!)(1!)}$$

$$= 50400 \text{ วิธี}$$

(4) การจัดกลุ่ม (Combination) : การจัดกลุ่มก็คือการจัดเรียงโดยไม่คำนึงถึงอันดับ เรามีวิธีการคิดคำนวณดังกฎต่อไปนี้

กฎที่ 5 สิ่งของที่ต่างกัน n สิ่ง จะนำมาจัดกลุ่มทีละ r สิ่ง โดยไม่แทนที่ จะได้จำนวนวิธีการจัดกลุ่มเป็น

$$C(n,r) = n!/(r!\,(n-r)!) \;, \quad r \leq n$$

จะเห็นได้ว่า (1) การจัดกลุ่มเหมือนกับการจัดเรียงอันดับสิ่งของ n สิ่ง ที่มี r สิ่งเหมือนกัน และอีก (n-r) สิ่ง เหมือนกัน นั่นเอง (ดูกฎที่ 4)

(2) การจัดกลุ่ม และการจัดเรียงอันดับมีความสัมพันธ์กันดังนี้

$$P(n,r) = r!C(n,r)$$

ตัวอย่าง เลือกกรรมการชุมนุมจิตวิทยา 4 คน จากผู้สมัคร 9 คน โดยการจับฉลากวิธีที่กรรมการจะได้รับเลือกมีเท่าใด ?

$$C(9,4) = 9!/(4!(9-4)!) = 126$$

หมายเหตุ (1) C(n, r) อาจจะเขียนได้เป็น $\binom{n}{r}$, ${}^{n}C_{r}$, nCr

(2) C(n, r) มักจะเรียกว่า**สัมประสิทธิ์ทวินาม** (Binomial Coefficients) เพราะมันปรากฏในการกระจาย $(a+b)^n$ ของทฤษฎีทวินาม (Binomial Theorem) นั่นคือ

$$(a+b)^n = \sum_{r=0}^{n} C(n, r)\, a^r\, b^{n-r}$$

(3) ถ้า n และ r เป็นเลขบวกจำนวนเต็มโดยที่ $0 \le r \le n$ แล้วเรามีทฤษฎีเกี่ยวกับสัมประสิทธิ์ทวินามที่น่าสนใจดังนี้

1. $\binom{n}{r} = \binom{n}{n-r}$

2. $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$

กฎที่ 6 สิ่งของที่เท่ากัน n สิ่ง จะนำมาจัดกลุ่มทีละ r สิ่ง โดยการแทนที่จะได้วิธีการจัดกลุ่มเป็น

$$C_1(n, r) = \frac{(n+r-1)!}{r!\,(n-1)!}\ ;\ r \ge 1$$

ตัวอย่าง มีคน 3 คน และมีงานอยู่ 2 งาน จะมีวิธีกำหนดงานให้คนทำได้กี่วิธี?

$$C_1(3, 2) = \frac{(3+2-1)!}{n!\,(3-1)!} = 6$$

ตัวอย่าง มีคน 2 คน แต่มีงานอยู่ 3 งาน จะมีวิธีกำหนดงานให้คนทำได้กี่วิธี?

$$C_1(2, 3) = \frac{(2+3-1)!}{3!\,(2-1)!} = 4$$

กฎที่ 7 สิ่งของที่ต่างกัน n สิ่ง ถ้าแบ่งออกเป็น k พวก (Cell or Subsets) โดยให้พวกที่ 1 มี n_1 สิ่ง, พวกที่ 2 มี n_2 สิ่ง,....,พวกที่ k มี n_k สิ่ง แล้วจำนวนวิธีจัดกลุ่มเป็น

$$C_2(n, n) = \frac{n!}{(n_1!)(n_2!)\ldots(n_k!)}\ ;\quad n_1 + n_2 + \ldots + n_k = n$$

ตัวอย่าง คนงาน 10 คน ถ้ากำหนดให้ไปทำงาน 3 ชนิด ชนิดแรกต้องการใช้คน 3 คน ชนิดที่สองใช้คน 5 คน และชนิดที่สามใช้ 2 คน จะมีวิธีกำหนดได้กี่วิธี?

$$C_2(10, 10) = 10!/(3!5!2!) = 2520 \text{ วิธี}$$

2.6 วิธีวัดความน่าจะเป็น (Methods of Measuring Probability)

จากนิยามที่ว่า ความน่าจะเป็นคือมาตรวัดชนิดหนึ่งที่ใช้วัดความไม่แน่นอนของการเกิดขึ้นของเหตุการณ์ นั่นคือ ใช้วัดว่าเหตุการณ์หนึ่งจะมีโอกาสเกิดขึ้นมากน้อยเพียงไร ผลที่วัดออกมาจะเป็นจำนวนเลข จำนวนเลขนั้นมักจะอยู่ในรูปทศนิยม, สัดส่วน, หรือเปอร์เซ็นต์ ในการกำหนดจำนวนเลขให้แก่เหตุการณ์หนึ่ง หรือในการวัดความน่าจะเป็นของเหตุการณ์หนึ่ง ๆ นั้นมีวิธีการกำหนดอยู่ 2 วิธีคือ

ก. วิธีปรนัย (Objective View) วิธีวัดแบบนี้จะได้ความน่าจะเป็นเชิงปรนัย (Objective Probability) ซึ่งอาศัยหลักเกณฑ์ดังนี้

(1) หลักความจริง หรือหลักเหตุผล (Axiomatic approach or Logical View) หรือหลักสมมาตร (Symmetry Principle) การวัดที่อาศัยเกณฑ์นี้พัฒนามาจากเกมส์พนันที่ยึดถือความจริง หรือเหตุผล และข้อสมมุติที่ว่า "ผลทดลองที่เป็นไปได้ทั้งหมดจากการทดลองจะไม่มีผลร่วมกัน แต่มีโอกาสจะเกิดขึ้นเท่า ๆ กัน (Mutually exclusive and Equally likely) ดังนั้นจำนวนเลขหรือความน่าจะเป็นที่จะกำหนดให้แก่แต่ละผลทดลองจึงเท่า ๆ กัน นั่นคือ ถ้ามีผลทดลองที่เป็นไปได้ทั้งหมดเท่ากับ n แล้ว เราจะกำหนดจำนวนเลข หรือความน่าจะเป็นให้แก่แต่ละผลทดลองเป็น 1/n เช่น ทอดลูกเต๋า ความน่าจะเป็นที่จะได้หน้าใดหน้าหนึ่งเท่ากันหมด คือ 1/6

สำหรับเหตุการณ์หนึ่งที่ประกอบด้วยผลทดลองที่เป็นไปได้จำนวนหนึ่ง เรามีวิธีการกำหนด ความน่าจะเป็นได้ดังนี้

นิยาม ถ้ากลุ่มผลทดลอง S ประกอบด้วยผลทดลองที่ไม่มีผลร่วมกันเลย แต่มีโอกาสเกิดขึ้นเท่า ๆ กัน เป็นจำนวน n(S) และถ้า E เป็นเหตุการณ์ที่สนใจ ซึ่งกำหนดขึ้นจาก S ประกอบด้วยผลทดลองเป็นจำนวน n(E) แล้ว ความน่าจะเป็นของเหตุการณ์ E, P(E), จะเป็นอัตราส่วนของ n(E) ต่อ n(S) นั่นคือ $P(E) = n(E)/n(S)$

ตัวอย่าง ในการทอดลูกเต๋า ถ้าสนใจหน้าที่ขึ้น เราจะได้จำนวนผลทดลองที่เป็นไปได้ทั้งหมด, n(S) เป็นจำนวน 6 และถ้าสนใจเหตุการณ์ที่ได้หน้าคู่ ซึ่งมีจำนวนผลทดลอง n(E) เป็น 3 ดังนั้นความน่าจะเป็นของเหตุการณ์ที่จะได้หน้าคู่, P(E) เป็น 3/6 หรือ 1/2

การกำหนดความน่าจะเป็นแบบนี้ John Maynard Keynes เป็นผู้กำหนดคนแรกในผลงานเรื่อง "A Treatise on Probability", (1921)

(2) หลักการทดลอง (Empirical Approach) หรือหลักความถี่สัมพัทธ์ (Relative Frequency Principle) กำหนดความน่าจะเป็นของเหตุการณ์ที่สนใจในเทอมของความถี่สัมพัทธ์ระยะยาวของการเกิดขึ้น (Long-run relative frequency of Occurrence) นั่นคือคิดเป็นอัตราส่วนของจำนวนครั้งที่เหตุการณ์เกิดขึ้นกับจำนวนครั้งทั้งหมดที่ทำการทดลอง และจะเรียกอัตราส่วน

นี้ว่า ความถี่สัมพัทธ์ของเหตุการณ์ ถ้าทำการทดลองซ้ำ ๆ กันเป็นจำนวนมากครั้งแล้ว เราเรียกอัตราส่วนนั้นว่า ความน่าจะเป็นของเหตุการณ์นั้น

เนื่องจากคำว่า ระยะยาว และจำนวนมากครั้งที่กล่าวมา เป็นคำที่เข้าใจยาก จึงหลีกเลี่ยงคำเหล่านี้ในการกำหนดความน่าจะเป็น แต่จะกำหนดว่าความน่าจะเป็นของเหตุการณ์หนึ่งเป็นขีดจำกัด (Limit) ของความถี่สัมพัทธ์ที่เหตุการณ์นั้นเกิดขึ้น ดังนิยามต่อไปนี้

นิยาม ในการทดลอง n ครั้ง ถ้าเหตุการณ์ที่สนใจเกิดขึ้น f ครั้ง ความถี่สัมพัทธ์ f/n (ในเมื่อ n ไม่จำกัดจำนวน) จะเข้าใกล้เลขจำนวนหนึ่ง เลขจำนวนนี้จะเรียกว่าขีดจำกัดของความถี่สัมพัทธ์ ดังนั้น ความน่าจะเป็นของเหตุการณ์ E กำหนดได้ว่าเป็นขีดจำกัดของความถี่สัมพัทธ์ของมันนั่นคือ

$$P(E) = \lim_{n \to \infty} f/n$$

นิยามเช่นนี้ ยุ่งยากในทางปฏิบัติ เพราะจำนวนครั้งที่ทำการทดลองนั้น เราทำได้เพียงจำกัดจำนวนครั้งเท่านั้น ไม่ใช่ไม่จำกัดจำนวนครั้ง แต่อย่างไรก็ตามจากกฎของธรรมชาติที่บอกให้เราทราบว่า "ในการทดลองต่าง ๆ ที่สามารถทำซ้ำ ๆ ได้ครั้งแล้วครั้งเล่า ภายใต้สภาวการณ์คงที่อย่างหนึ่งนั้น จะมีคุณลักษณะที่น่าสนใจอย่างหนึ่ง คือ ถ้าทำการทดลองซ้ำ ๆ เป็นจำนวนครั้งไม่จำกัดความถี่สัมพัทธ์หรือสัดส่วนของครั้งในเหตุการณ์ใด ๆ ที่เกิดขึ้นจะคงที่มากขึ้น ๆ ขณะที่จำนวนครั้งของการทดลองได้เพิ่มขึ้น ๆ และแนวโน้มของสัดส่วนนี้ จะมุ่งเข้าหาค่าคงที่ค่าหนึ่ง คุณลักษณะเช่นนี้ได้ชื่อว่า ความปกติเชิงสถิติ (Statistical Regularity) หรือความคงที่ของความถี่สัมพัทธ์ (Stability of Relative Frequencies) ดังนั้น นิยามที่ใช้วัดความน่าจะเป็นโดยอาศัยหลักการทดลองที่นิยมกันมากก็คือ

นิยาม ในการทดลองเชิงสุ่มที่ทำการทดลองจำนวน n ครั้ง เหตุการณ์ที่สนใจ E เกิดขึ้น f ครั้ง สัดส่วน f/n จะมีแนวโน้มจะคงที่ (Stabilize) ณ ค่า P(E) ในเมื่อ n เพิ่มขึ้นไม่จำกัดจำนวน และสัดส่วน f/n นี้ จะเรียกว่า ความน่าจะเป็นของเหตุการณ์ E นั้น นั่นคือ

$$P(E) = f/n$$

เราเห็นได้ว่าจำนวนครั้ง f ที่เกิดขึ้นของเหตุการณ์จะผันแปรจากค่าต่ำสุด (นั่นคือเหตุการณ์ไม่เกิดขึ้นสักครั้ง) ถึงค่าสูงสุด (นั่นคือเหตุการณ์จะเกิดขึ้นทุกครั้งที่ทำการทดลอง) ดังนั้น ความว่าจะเป็นของเหตุการณ์ E จะผันแปรจากค่าต่ำสุด ถึงสูงสุด นั่นคือ

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

สำหรับการทดลองที่จำกัดจำนวนครั้ง ความถี่สัมพัทธ์ f/n ที่สังเกตได้จะเป็นแต่เพียงค่าประมาณ (Estimate) ของค่าจริง ๆ ของ P(E) เท่านั้น อย่างไรก็ตามถ้าจำนวนครั้งของการทดลองมากพอ เราก็หวังอย่างมีเหตุผลว่า ความถี่สัมพัทธ์ที่สังเกตได้จะใกล้เคียงกับความน่าจะเป็นจริง ๆ ของเหตุการณ์ E นั้น เช่น เมื่อสงครามโลกครั้งที่สอง J.E. Kerrick ได้ทอดเหรียญ 10,000 ครั้ง

ปรากฏว่า ขึ้นหัว 5067 ครั้ง ดังนั้น 5067/10000 เป็นค่าประมาณของค่าจริงคือ 0.50 ซึ่งเป็นความน่าจะเป็นจริงของเหตุการณ์ที่สตางค์จะขึ้นหัว

การจัดความน่าจะเป็นโดยวิธี John Venn เป็นผู้สนใจมาก่อน ต่อมาก็มี Von Mises และ Reichenbach

ข. วิธีอัตนัย (Subjective View) โดยวิธีนี้เราจะได้ความน่าจะเป็นเชิงจิตวิสัย (Subjective Probability)

ในการกำหนดความน่าจะเป็นเชิงปรนัยที่กล่าวมาแล้วว่าเป็นความถี่สัมพัทธ์นั้นจะมีเหตุผลก็ต่อเมื่อการทดลองสามารถทำซ้ำ ๆ ได้ แต่บางครั้งเราก็ทำไม่ได้ ดังนั้น จึงจำเป็นต้องหาวิธีวัดความน่าจะเป็นของเหตุการณ์ที่สนใจด้วยวิธีอื่น นั่นคือ ได้ใช้วิธีอัตนัย ซึ่งวิธีนี้ยึดตัวบุคคลเป็นหลัก และได้ใช้ระดับความเชื่ออย่างมีเหตุผล (Degree of Rational Belief) เป็นประโยชน์ในการกำหนดความน่าจะเป็น ปัจจุบันการกำหนดความน่าจะเป็นวิธีนี้กลายเป็นหลักสำคัญในทฤษฎีการตัดสินใจ

ตามปกติสภาวการณ์ของการตัดสินใจที่ประสบอยู่มาก ไม่ได้เกิดขึ้นมาในอดีตและบางครั้งก็ไม่เกิดขึ้นอีกในแบบเดิม ดังนั้น จึงไม่มีความถี่สัมพัทธ์ของการเกิดขึ้นสำหรับเหตุการณ์ต่าง ๆ ที่สนใจหรือไม่มีพื้นฐานในหลักเหตุผลอย่างแจ่มชัดเพื่อที่จะกำหนดความน่าจะเป็นนั่นก็คือ ทำให้เราไม่สามารถกำหนดในเชิงปรนัยได้ ตัวอย่างเช่น ในการแข่งม้า อาจจะมีม้าเข้าแข่ง 8 ตัวม้าใดจะเข้าหลักชัยที่หนึ่งนั้น เป็นผลจากการทดลองเชิงสุ่ม ซึ่งบอกไม่ได้เลยว่าแต่ละตัวนั้นมีโอกาสที่จะเข้าหลักชัย เป็นที่หนึ่งได้เท่า ๆ กัน (คือเท่ากับ 1/8) ทั้งนี้ เพราะทำให้เกิดขึ้นซ้ำ ๆ ในสภาวะการณ์เดียวกันไม่ได้เราจึงวัดความน่าจะเป็นแบบวิธีปรนัยไม่ได้ ดังนั้น วิธีที่เราจะวัดความน่าจะเป็นได้ก็คือ ใช้วิธีประมาณโดยการชั่งใจของเราเองในเชิงจิตวิสัย หรืออัตนัย ซึ่งใช้ความรู้ที่ว่า ก่อนแข่งม้าเขามักจะเอาม้ามาเดินผ่านอัฒจรรย์ให้คนได้ดูลักษณะท่าทางของม้า และอาจมีสถิติการวิ่งเก่า ๆ พิมพ์แจก นั่นก็คือทำให้เราสามารถกำหนดความน่าจะเป็นคร่าว ๆ ว่าความน่าจะเป็นที่ม้าตัวหนึ่งตัวใดจะเข้าหลักชัยเป็นเท่าใด

การกำหนดความน่าจะเป็นเช่นนี้เป็นการกำหนดเชิงจิตวิสัย คือ ใช้ความรู้สึกของตนเองเป็นเครื่องกำหนด การกำหนดแบบนี้จะถูกต้องตามความน่าจะเป็นที่แท้จริงหรือใกล้เคียงเท่าไรก็ไม่สามารถพิสูจน์ได้ในเฉพาะกรณี แต่จากประสพการณ์ที่ผ่านมาได้แสดงให้เห็นว่าการใช้ความน่าจะเป็นเชิงจิตวิสัยมาช่วยประกอบในการตัดสินใจ จะได้ผลประโยชน์บ้าง ซึ่งดีกว่าที่จะไม่ใช้ความน่าจะเป็นเสียเลย

การวัดหรือกำหนดในเชิงอัตนัยนี้ James Bernoulli ได้สนใจและพิมพ์หนังสือชื่อ Ars Conjectandi (the Art of Guessing) ซึ่งได้กำหนดความน่าจะเป็นในเทอมของดีกรีของความเชื่อมั่นและ August De Morgan ได้กำหนดความน่าจะเป็นในเทอมของความเชื่อ ต่อมา L.J.

Savage R.Schlaifer และ H.Raiffa ได้นำการกำหนดความน่าจะเป็นเชิงอัตนัยไปใช้ในทฤษฎีการตัดสินใจ

2.7 ฟังก์ชัน และ สัจจพจน์ความน่าจะเป็น (Probability Functions and Axioms)

นิยาม ให้ S เป็นกลุ่มผลทดลองของการทดลองเชิงสุ่ม และให้ $A_1, A_2, A_3, \ldots\ldots$ เป็นเหตุการณ์ใน S หรือ $A_i \subset S$ (i= 1,2,3,..) ฟังก์ชัน P ซึ่งเป็นฟังก์ชันที่กำหนดจำนวนเลขจริง $P(A_i)$ ให้แก่แต่ละเหตุการณ์ A_i ที่อยู่ใน S นั้นจะเรียกว่า ฟังก์ชันน่าจะเป็น ถ้าค่า $P(A_i)$ ของฟังก์ชันสอดคล้องกับสัจจพจน์ต่อไปนี้

สัจจพจน์ 1. $0 \leq P(A_i) \leq 1$; i = 1, 2, 3,......

สัจจพจน์ 2. $P(S) = 1$

สัจจพจน์ 3. $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ ในเมื่อ $A_i \cap A_j = \phi, i \neq j$

$P(A_i)$ เรียกว่า ความน่าจะเป็นของเหตุการณ์ A_i และ คู่ (S,P) เรียกว่าตัวแบบน่าจะเป็น (Probability Model or Space)

บทแทรกของสัจจพจน์ 3 ถ้า s_i เป็นผลทดลองที่เป็นไปได้ที่อยู่ในเหตุการณ์ A_i แล้ว

$$P(A_j) = \sum_{s_i \in A_i} P(s_i)$$

ตัวอย่าง ทอดลูกเต๋าครั้งหนึ่ง และสังเกตหน้าที่ขึ้น

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

ถ้าความน่าจะเป็นที่จะกำหนดให้แก่แต่ละผลทดลองตามหลักความจริงแล้ว เราจะได้ฟังก์ชันน่าจะเป็นดังนี้

เหตุการณ์ A_i	1	2	3	4	5	6
ความน่าจะเป็น $P(A_i)$	1/6	1/6	1/6	1/6	1/6	1/6

เราสามารถตรวจสอบได้ว่าตารางนี้เป็นฟังก์ชันน่าจะเป็น เพราะ

(1) $P(A_i)$ เท่ากับ 1/6 ซึ่งอยู่ระหว่าง 0 กับ 1

(2) $P(S) = 1$

(3) กำหนด A_1, A_2 โดยที่ $A_1 \cap A_2 = \emptyset$ จะได้ว่า

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \text{ นั่นคือ } 2/6 = 1/6 + 1/6$$

ตัวอย่าง นักธุรกิจผู้มั่งคั่งคนหนึ่งมีหุ้นอยู่ 3 ชนิด สาธารณูปโภค, อุตสาหกรรม และธนาคาร ซึ่งคิดเป็นเปอร์เซ็นต์ตามลำดับ ดังนี้ 20, 30, และ 50 ถ้าเลือกหุ้นมาหุ้นหนึ่ง (แต่ละหุ้นที่จะถูกเลือกเท่า ๆ กัน) แล้วแบบของหุ้นที่จะเลือกจะเป็นแบบใดแบบหนึ่งใน A = {สาธารณูปโภค, อุตสาหกรรม, ธนาคาร}

ดังนั้น P(สาธารณูปโภค) = 20/100 = .20

P (อุตสาหกรรม) = 20/100 = .30

P (ธนาคาร) = 50/100 = .50

ข้อสังเกต ฟังก์ชั่น (Function) เป็นคำที่ใช้กันบ่อย ๆ ในทางคณิตศาสตร์ ความน่าจะเป็นและสถิติ สามารถอธิบายได้ดังนี้

กำหนดให้ A และ B เป็นเซท และกฎของการสัมนัย (Rule of Correspondence) ที่กำหนดความสัมพันธ์ระหว่างสมาชิกของ A กับสมาชิกของ B ในลักษณะที่ว่าทุก ๆ สมาชิก x ของ A จะมีสมาชิก ตัวหนึ่งและตัวเดียว (Unique) ของ B ที่สมนัยกับ x แล้วกฎนี้จะระบุเซทของคู่ลำดับ (Ordered pairs), f และเซท f นี้ เรียกว่าฟังก์ชั่นจาก A ไป B (function from A to B)

ฟังชั่น f สามารถเขียนได้เป็น $f = \{(x, y)/$สำหรับทุก $x \in A$ จะมี $y \in B$ ตัวหนึ่งตัวเดียวเท่านั้น$\}$

มีสิ่งที่น่าสนใจจากนิยามของฟังก์ชั่นดังนี้

(1) ฟังก์ชั่นเป็นเซท

(2) ฟังก์ชั่นจะแทนด้วยอักษรโรมัน เช่น f,g,h,F,G,.....เป็นต้น

(3) สมาชิก y ในเซท B อาจเขียนได้ f(x) ในเมื่อ x เป็นสมาชิกหนึ่งใน A ดังนั้น f(x) และ y จึงแทนสิ่งเดียวกัน

(4) ฟังก์ชั่นจาก A ไป B จะให้เซทของคู่ลำดับในรูป (x,y) หรือ (x, f(x)) และฟังก์ชั่นจาก B ไป A จะให้เซทคู่ลำดับในรูป (y, x) หรือ (y, f(y))

(5) ถ้า f เป็นฟังก์ชั่นจาก A ไป B แล้วเซท A จะเรียกว่าโดเมน (Domain) ของฟังก์ชั่น และเซท B เรียกว่า พิสัย (Range) ของฟังก์ชั่น กระบวนการที่สร้างความสมนัย หรือคู่ลำดับเรียกว่าการฉาย (mapping) เซท A ที่ฉาย (map) เข้าไปในเซท B จะแทนด้วย $A \longrightarrow B$

ฟังก์ชั่นที่มีพิสัยเป็นเลขจำนวนจริง (Real numbers) จะเรียกว่าฟังก์ชั่นค่าจริง (Real-valued function) ในทฤษฎีน่าจะเป็นจะใช้ฟังก์ชั่นค่าจริงนี้

ในการกล่าวถึงฟังก์ชั่น เรามักจะสรุปได้ดังนี้

(1) ฟังชั่นมักจะระบุโดยการบรรยายถึงกฎของการสมนัยเท่านั้น ซึ่งบ่อยครั้งจะอยู่ในรูปของสมการ

(2) โดเมนและพิสัยของฟังก์ชั่นมักจะสมมุติว่าเป็นเซทของเลขจำนวนจริง

(3) x จะใช้แทนสมาชิกในโดเมนของฟังก์ชั่น

(4) y = f(x) ใช้แทนสมาชิกในเซทพิสัย (Range set)

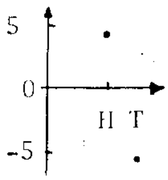
ฟังก์ชั่นสามารถแสดงได้หลายวิธี ดังนี้

(1) เซทของคู่ลำดับ $f = \{(H,5),(T,-5)\}$
(2) กราฟ
(3) ตาราง
(4) แผนภาพ
(5) สูตร หรือสมการ
(6) เครื่องจักร

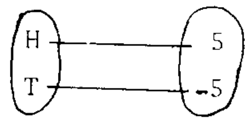
(3) ตาราง

ผลการทดลอง	เงินที่ได้รับ
H	5
T	-5

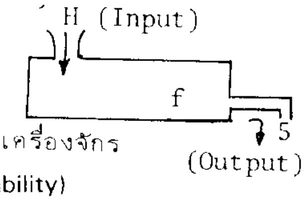
(2) กราฟ



(4) แผนภาพ



(5) สูตรหรือ สมการ $f(H) = 5$
$f(T) = -5$



(6) เครื่องจักร

2.8 ความน่าจะเป็นร่วมและทางเดียว (Joint and Marginal Probability)

ในกลุ่มผลทดลอง S ซึ่งประกอบด้วย n ผลทดลองหรือจุด และแต่ละจุดมีความน่าจะเป็น 1/n ถ้ามีเหตุการณ์ $A_1, A_2, \ldots, A_m$ ประกอบเป็นส่วนแบ่งของ S และเหตุการณ์ $B_1, B_2, \ldots, B_n$ ก็เป็นส่วนแบ่งของ S อีก แล้วเราสามารถสร้างกลุ่มผลทดลองเป็นตาราง 2 ทาง (Two-way table) ได้ดังนี้

	B	B_1	B_2	B_3	B_4	$\cdots$	B_n	
A	A_1	n_{11}	n_{12}	n_{13}	n_{14}	$\cdots$	n_{1n}	n_1
	A_2	n_{21}	n_{22}	n_{23}	n_{24}	$\cdots$	n_{2n}	n_2
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	A_m	n_{m1}	n_{m2}	n_{m3}	n_{m4}		n_{mn}	n_m
		$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot 3}$	$n_{\cdot 4}$		$n_{\cdot n}$	n

ในเมื่อ n_{ij} $(i=1,2,\ldots,m;\ j=1,3,\ldots,n)$ เป็นจำนวนจุดใน n จุดที่มีลักษณะ (attribute) ของ A และ B_j เช่น n_{11} มีลักษณะของ A_1 และ B_1, n_{12} มีลักษณะของ A_1 และ B_1, n_{12} มีลักษณะของ A_1 และ B_2

ผลรวมของทุก ๆ n_{ij} จะเท่ากับ n นั่นคือ $\sum_i^m \sum_j^n n_{ij} = n$

ความน่าจะเป็นของเหตุการณ์ A_i และ B_j หรือ $P(A_i \cap B_j)$ นั้นจะเท่ากับ n_{ij}/n ซึ่งจะเรียกว่า ความน่าจะเป็นร่วม (joint Probability) ของเหตุการณ์ A_i และ B_j

ถ้าสมมุติว่าเราสนใจเพียงลักษณะเดียว เช่น A (ไม่สนใจลักษณะ B) แล้วความน่าจะเป็นของ A_2 (ลักษณะหนึ่งของ A) หรือ $P(A_2)$ จะเป็นดังนี้

$$P(A_2) = (n_{21}+n_{22}+\cdots+n_{2n})/n = \sum_j^n n_{2j}/n$$

ซึ่งเรียกว่า ความน่าจะเป็นทางเดียว (Marginal Probability)

โดยทั่วไปเราจะได้ว่า $P(A_i) = \sum_{j=1}^{n} n_{ij}/n = \sum_{j=1}^{n} P(A_i \cap B_j)$

และในทำนองเดียวกัน ความน่าจะเป็นทางเดียวของ B_j คือ

$$P(B_j) = \sum_{i=1}^{m} P(A_i \cap B_j)$$

ถ้าดูในเทอมของเซทก็จะเห็นว่าเราแบ่งกลุ่มผลทดลอง S ออกเป็น mn เซทย่อยซึ่งเราจะได้

$$A_i = (A_i \cap B_1) \cup (A_i \cap B_2) \cup \ldots\ldots, \cup (A_i \cap B_n)$$

ดังนั้น $P(A_i) = P(A_i \cap B_1) + P(A_i \cap B_2) + \ldots + P(A_i \cap B_n)$

เพราะ $(A_i \cap B_j) \cap (A_i \cap B_k) = \phi$; $j, k = 1, 2, \ldots, n \quad j \neq k$

นั่นคือ $\quad P(A_i) = \sum_{j}^{n} P(A_i \cap B_j) = \sum_{j}^{n} n_{ij}/n$

ตัวอย่าง นักวิจัยได้สัมภาษณ์หัวหน้าครอบครัวเกี่ยวกับรายได้ต่อปี และจำนวนรถที่ครอบครองอยู่ เป็นจำนวนครอบครัว 600 ราย ได้ผลดังนี้

รายได้ต่อปี (บาท)	-15,000	15,000-37,500	37,500-50,000	$50,000^+$
จำนวนรถ 2^+	1	10	30	180
จำนวนรถ 1	21	100	200	30
จำนวนรถ 0	10	8	6	4

ให้ C_1, C_2 และ C_3 เป็นจำนวนรถที่มีอยู่ในครอบครัวเป็นจำนวน 2 คัน หรือ มากกว่า, เพียงคันเดียว, และไม่มีรถเลย ตามลำดับ

และให้ R_1 , R_2, R_3 และ R_4 เป็นครอบครัวที่มีรายได้น้อยกว่า 15,000; 15,000 ถึง 37,500 ; 37,500 - 50,000 ; และ 50,000 บาทขึ้นไป ตามลำดับ

เราสามารถประมาณความน่าจะเป็นของเหตุการณ์ต่าง ๆ ได้ดังนี้

$P(C_2 \cap R_3) = 200/600 = 2/3$

$P(R_4 \cap C_1) = 180/600 = 3/10$

$P(C_1) = (1 + 10 + 20 + 80)/600 = 211/600$

$P(C_2) = (21 + 100 + 200 + 30)/600 = 361/600$

$P(C_3) = (10 + 8 + 6 + 4)/600 = 28/600$

$P(R_1) = (1 + 21 + 10)/600 = 32/600$

$P(R_2) = (10 + 100 + 8)/600 = 118/600$

$P(R_3) = (30 + 200 + 6)/600 = 236/600$

$P(R_4) = (180 + 30 + 4)/600 = 214/600$

2.9 ความน่าจะเป็นเงื่อนไข (Conditional Probability)

ตามปกติเราจะหาความน่าจะเป็นของเหตุการณ์ใด เราก็เอาเหตุการณ์นั้นไปเทียบกับกลุ่มผลทดลอง (Sample Space) ของมัน ดังตัวอย่าง

ตัวอย่าง ทอดลูกเต๋า 2 ครั้ง ความน่าจะเป็นที่จะได้ผลรวมเป็น 7 จะเท่ากับ 6/36 ทั้งนี้เพราะว่า

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

$n(S) = 36$

ผลรวมที่จะเป็น 7 $E = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

$n(E) = 6$

แต่ E สามารถเขียนได้เป็น $E \cap S$ ดังนั้น

$n(E) = n(E \cap S) = 6$

เพราะฉะนั้นความน่าจะเป็นของ E หรือ $P(E) = \dfrac{n(E \cap S)}{n(S)} = 6/36$

ความน่าจะเป็นที่กำหนดโดยการนำไปเทียบกับกลุ่มผลทดลองนี้บางทีเรียกว่า ความน่าจะเป็นไม่มีเงื่อนไข (Unconditional Probability)

บางครั้งเราต้องการจะนำเหตุการณ์ที่สนใจไปเทียบกับเหตุการณ์อื่น เราจะเรียกความน่าจะเป็นนั้นว่า ความน่าจะเป็นเงื่อนไข (Conditional Probability) เหตุการณ์อื่นนี้จะมีผลทำให้เหตุการณ์ที่เราสนใจมีความน่าจะเป็นเปลี่ยนไปจากเดิม คือเปลี่ยนไปจากที่เราเทียบกับกลุ่มผลทดลอง เหตุการณ์อื่นที่เรานำมาเปรียบเทียบนี้จะทำหน้าที่เหมือนกลุ่มผลทดลองซึ่งเรียกว่ากลุ่มผลทดลองทดแทน (Reduced Sample Space)

ตัวอย่างเช่น ถ้าทราบว่าผลรวมของลูกเต๋าที่เป็น 7 นั้น มีหน้า 5 อยู่ด้วยแล้ว ความน่าจะเป็นของผลรวมเป็น 7 จะเปลี่ยนไปอย่างไร เมื่อให้ E_1 เป็นเหตุการณ์ที่เกิดหน้า 5 นั่นคือ

$E_1 = \{(5,1)(5,2), \ldots\ldots, (1,5)\}$

ดังนั้น P(E เทียบกับ E_1) $= P(E/E_1) = n(E \cap E_1)/n(E_1)$

$= 2/11$

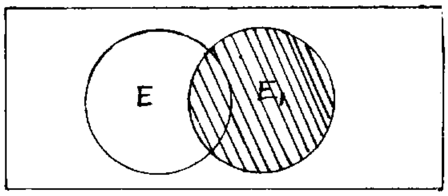
เพราะว่า $E \cap E_1 = \{(5,2), (2,5)\}$, $n(E \cap E_1) = 2$, $n(E) = 11$

เราจะเห็นได้ว่าความน่าจะเป็นเปลี่ยนจาก 1/6 เป็น 2/11

เมื่อเหตุการณ์ E_1 ทำให้ความน่าจะเป็นของ E เปลี่ยนไปเช่นนี้ เราจะพูดว่า E ขึ้นอยู่กับ E_1 หรือ E และ E_1 ขึ้นอยู่แก่กัน (Statistically Dependent) ต่อไปจะให้นิยามของความน่าจะเป็นเงื่อนไข

นิยาม : ให้ S เป็นกลุ่มผลทดลอง และ E กับ E_1 เป็นเหตุการณ์ที่กำหนดใน S ความน่าจะเป็นของ E กำหนดว่า E_1 เกิดขึ้นแล้วนั้น เราเรียกว่า ความน่าจะเป็นเงื่อนไขของ E จะเขียนได้เป็น $P(E/E_1)$ ซึ่งกำหนดไว้ดังนี้

$$P(E/E_1) = P(E \cap E_1)/P(E_1); \qquad P(E_1) \neq 0$$



กลุ่มผลทดลองเดิม
(OLD Sample Space)
กลุ่มผลทดลองใหม่
(Reduced Sample Space)

เราจะเห็นว่าเมื่อ E_1 เกิดขึ้นแล้วนั้นทำให้สมาชิกของ S เกิดขึ้นไม่ทั้งหมด ส่วนที่เกิดขึ้นจะเป็นของ E_1 ทั้งหมดนั่นคือ $E_1 \subset S$ เราจะได้อีกว่าสมาชิกของE จะเกิดขึ้นบางส่วน และส่วนที่เกิดขึ้นก็จะเป็นของ E_1 ด้วย นั่นคือ ส่วนที่เกิดจะเป็นของ $E \cap E_1$ นั่นเอง

ดังนั้น ความน่าจะเป็นของ E ที่เกิดขึ้นโดยกำหนด E_1 เกิดขึ้นแล้ว $P(E/E_1)$ เราจะหาได้ดังนี้

$$P(E/E_1) = n(E \cap E_1)/n(E_1) = \frac{n(E \cap E_1)/n(S)}{n(E_1)/n(S)}$$
$$= P(E \cap E_1)/P(E_1)$$

ในทำนองเดียวกัน เราจะได้ว่า

$$P(E_1/E) = P(E_1 \cap E)/P(E)$$

เราสามารถแสดงได้ว่า $P(E/E_1)$ นี้สอดคล้องกับสัจจพจน์ของความน่าจะเป็นด้วยนั่นคือ

(1) $0 \leq P(E/E_1) \leq 1$

(2) $P(S/E_1) = 1$

(3) $P(A_1 \cup A_2 \cup \ldots /E_1) = \sum P(A_i/E_1)$ ถ้า $A_i \cap A_j = \emptyset$; $i \neq j$

ตัวอย่าง ในสำนักงานมีเครื่องคำนวณอยู่ 100 เครื่อง ซึ่งมีทั้งเครื่องไฟฟ้า (E) และเครื่องจักรกล (M) และมีทั้งเครื่องใหม่ (N) และใช้แล้ว (U) จำนวนของแต่ละอย่างแสดงได้ดังตารางต่อไปนี้

	E	M	
N	40	30	70
U	20	10	30
	60	40	100

ถ้าพนักงานในสำนักงานคนหนึ่งต้องการใช้เครื่อง จึงไปหยิบเครื่องคำนวณมาเครื่องหนึ่งแบบสุ่ม และพบว่าเป็นเครื่องใหม่ จงหาโอกาสที่จะเป็นเครื่องไฟฟ้า

สิ่งที่เราต้องการทราบ คือ P(E/N) แต่เราทราบว่า N ซึ่งเป็นกลุ่มผลทดลองทดแทนมีจำนวน 70 เราจึงได้ว่า

$$P(E/N) = n(E \cap N)/n(N) = 40/70 = 4/7$$

แต่จากนิยาม

$$P(E/N) = P(E \cap N)/P(N) = (40/100)/(70/100) = 4/7$$

ซึ่งเท่ากับที่คำนวณมาแล้วนั่นเอง

ตัวอย่าง สมมุติว่าโอกาสที่ส่วนประกอบอย่างหนึ่งในระบบมีอายุใช้งานอย่างน้อย x ชั่วโมง เป็น e^{-7x} โอกาสที่ส่วนประกอบนี้จะมีอายุใช้งานเพิ่มขึ้นอีก y ชั่วโมง กำหนดว่ามันทำงานมาแล้ว z ชั่วโมง จะเป็นเท่าใด?

ให้ B เป็นเหตุการณ์ที่ส่วนประกอบในระบบมีอายุใช้งานอย่างน้อย Z ชั่วโมง ;

$B = \{x \in R / x \geq z$, R เป็นเซ็ทของเลขจำนวนจริง ๆ$\}$

และให้ A เป็นเหตุการณ์ที่ส่วนประกอบในระบบจะมีอายุใช้งานอย่างน้อย z + y ชั่วโมง ;

$A = \{x \in R / x \geq z + y\}$

แต่ $A \cap B = A$ ดังนั้น $P(A/B) = P(A \cap B)/P(B) = P(A)/P(B)$

$$= e^{-7(Z+y)} / e^{-7Z} = e^{-7y}$$

2.10 กฎของความน่าจะเป็น (Fundamental Probability Rules)

ในการกำหนดความน่าจะเป็นให้แก่เหตุการณ์หรือผลรวมของเหตุการณ์เรามีกฎหรือทฤษฎีซึ่งช่วยกำหนดความน่าจะเป็นของเหตุการณ์เหล่านั้นได้ ดังนี้

1. **กฎเติมเต็ม** (Complementation Rule) ถ้าเหตุการณ์ E และสวนเติมเต็ม ประกอบเป็นส่วนแบ่งของกลุ่มผลทดลอง S แล้ว

$$P(\bar{E}) = 1 - P(E) \; ; \quad P(E) = 1 - P(\bar{E})$$

ตัวอย่าง ปัญหาเกี่ยวกับวันเกิด (Classical Birthday Problem) นักศึกษาในห้องหนึ่งมี n คน โอกาสที่นักศึกษาอย่างน้อย 2 คน มีวันเกิดตรงกันเป็นเท่าใด? (สมมุติ 1 ปี มี 365 วัน)

ให้ E เป็นเหตุการณ์ที่นักศึกษาอย่างน้อย 2 คน มีวันเกิดร่วมกัน

$\bar{E}$ เป็นเหตุการณ์ที่นักศึกษาไม่มีวันเกิดร่วมกัน

เราจะเห็นได้ว่าวิธีหรือผลทดลองซึ่งเป็นไปได้ที่นักศึกษา n คน จะเกิดใน 365 วัน มี $365(365)\ldots\ldots(365) = 365^n$ วิธี

ส่วนวิธีที่นักศึกษาจะไม่มีวันเกิดร่วมกันมี 365 (364)...... [365 - (n - 1)] = 365 Pn วิธี

ดังนั้นเรากำหนดความน่าจะเป็นให้แก่เหตุการณ์ $\bar{E}$ ได้เป็น

$$P(\bar{E}) = {}_{365}P_n/365^n$$

นั่นคือ

$$P(E) = 1 - P(\bar{E}) = 1 - {}_{365}P_n/365^n$$

สำหรับ $P(E) = {}_{365}P_n/365^n$

$$= (365/365)(364/365)\ldots(365-(n-1))/365$$

$$= (1-1/365)(1-2/365)\ldots(1-(n-1)/365)$$

จากทฤษฎีทางคณิตศาสตร์ที่ว่า $\ln(1-x) \cong -x$

ดังนั้น $\ln P(E) \cong \sum_i (-i/365)$

นั่นคือ

$$P(E) \cong \exp(\sum_i (-i/365))$$

ถ้า n=20 เราจะได้ว่า $P(\bar{E}) = 0.412$

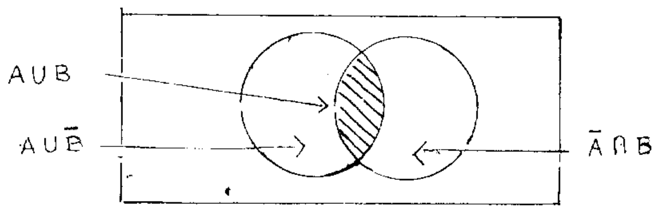
2. **กฎผลบวก** (Addition Rule) ถ้า A และ B เป็นสองเหตุการณ์ใด ๆ ใน S แล้ว

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

บทแทรก (1) ถ้า $A \cap B = \phi$ แล้ว $P(A \cup B) = P(A) + P(B)$

(2) ถ้า $A \cap B \neq \phi$ แล้ว $P(A) = P(A \cap B) + P(A \cap \bar{B})$

และ $P(\bar{B}) = P(A \cap B) + P(\bar{A} \cap B)$



(3) ถ้า A, B และ C เป็นสามเหตุการณ์ใดใน S แล้ว $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)$

(4) กฎผลบวกทั่วไป

ก) ถ้า $A_1, A_2, \ldots\ldots\ldots, A_k$ เป็น k เหตุการณ์ใด ๆ ใน S แล้ว

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

ข) ถ้า $A_i \cap A_j = \emptyset$; $i \neq j$ แล้ว

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

ตัวอย่าง แพทย์ผู้เชี่ยวชาญด้านมะเร็งได้รวบรวมข้อมูลเกี่ยวกับการตรวจคนไข้ไว้ดังนี้

ผู้ป่วยรู้สึกเป็นมะเร็ง และจากการตรวจก็พบว่าจริง 5%

ผู้ป่วยรู้สึกเป็นมะเร็ง แต่จากการตรวจก็ไม่พบ 45%

ผู้ป่วยไม่รู้สึกเป็นมะเร็ง แต่จากการตรวจก็พบว่าเป็น 10%

ผู้ป่วยไม่รู้สึกเป็นมะเร็ง และจากการตรวจก็ไม่พบว่าเป็น 40%

จงหาความน่าจะเป็นที่ (1) ผู้ป่วยรู้สึกว่าเป็นมะเร็ง

(2) แพทย์ตรวจพบว่าเป็นมะเร็ง

(3) แพทย์ตรวจว่าเป็นมะเร็ง **หรือ** ผู้ป่วยรู้สึกว่าเป็น

ให้ A เป็นเหตุการณ์ที่ผู้ป่วยรู้สึกว่าเป็นมะเร็ง

B เป็นเหตุการณ์ที่แพทย์ตรวจพบว่าเป็นมะเร็ง

	A	$\bar{A}$
B	.05	.10
$\bar{B}$	.45	.40

จากข้อมูลได้ว่า $P(A \cap B) = .05$, $P(A \cap B) = .45$

$P(A \cap \bar{B}) = .10$, $P(\bar{A} \cap \bar{B}) = .40$

ดังนั้น (1) $P(A) = P(A \cap B) + P(A \cap \bar{B}) = .05 + .45$

$= .50$

(2) $P(B) = P(A \cap B) + P(\bar{A} \cap B) = .05 + .10$

$= .15$

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

$= .50 + .15 - 05 = .60$

3. **กฎการคูณ** (Multiplication Rule) จากนิยามความน่าจะเป็นเงื่อนไข $P(E/E_1) = P(E \cap E_1) / P(E_1)$ หรือ $P(E_1 / E) = P(E_1 \cap E) / P(E)$ เราจะได้กฎการคูณดังนี้

ถ้า A และ B เป็นเหตุการณ์ในกลุ่มผลทดลอง S แล้ว

$$P(A \cap B) = \begin{cases} P(A/B)\,P(B) \\ P(B/A)\,P(A) \end{cases}$$

บทแทรก กฎการคูณทั่วไป : ถ้า $A_1, A_2, \ldots\ldots, A_n$ เป็นเหตุการณ์ใน S แล้ว

$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2/A_1) \ldots P(A_n/A_1A_2 \ldots A_{n-1})$

ตัวอย่าง ลังใบหนึ่งมีของอยู่ 10 ชิ้น ไม่ดี 4 ชิ้น ถ้าหยิบของแบบสุ่มมา 1 ชิ้น จงหาโอกาสของเหตุการณ์ต่อไปนี้

(1) ชิ้นต่อไปที่สุ่มมาอีกจะเป็นของไม่ดี เมื่อชิ้นแรกไม่ดี

(2) ชิ้นที่สามจะเป็นของไม่ดี (สองชิ้นแรกไม่ดี)

(3) ชิ้นที่สี่จะเป็นของดี โดยที่สามชิ้นแรกจะเป็นของไม่ดี

(4) ดีทั้ง 4 ชิ้น จากการหยิบ 4 ครั้ง

(5) 2 ชิ้นแรกดี แต่ 2 ชิ้นหลังไม่ดี

(6) 2 ชิ้นแรกไม่ดี แต่ 2 ชิ้นหลังดี

ให้ D และ G เป็นเหตุการณ์ที่หยิบของไม่ดี และดีตามลำดับ แล้วเราจะได้ว่า

(1) $P(D_1D_2) = P(D_1)P(D_2/D_1) = (4/10)(3/9) = 2/15$

(2) $P(D_1D_2D_3) = P(D_1)P(D_2/D_1)P(D_3/D_1D_2) = (4/10)(3/9)(2/8) = 1/30$

(3) $P(D_1D_2D_3G_4) = P(D_1)P(D_2/D_1)P(D_3/D_1D_2)P(G_4/D_1D_2D_3) = (4/10)(3/9)(2/8)(6/7) = 1/35$

(4) $P(G_1G_2G_3G_4) = (6/10)(5/9)(4/8)(3/7) = 1/14$

(5) $P(G_1G_2D_3D_4) = (6/10)(5/9)(4/8)(3/7) = 1/14$

(6) $P(D_1D_2G_3G_4) = (4/10)(3/9)(6/8)(5/7) = 1/14$

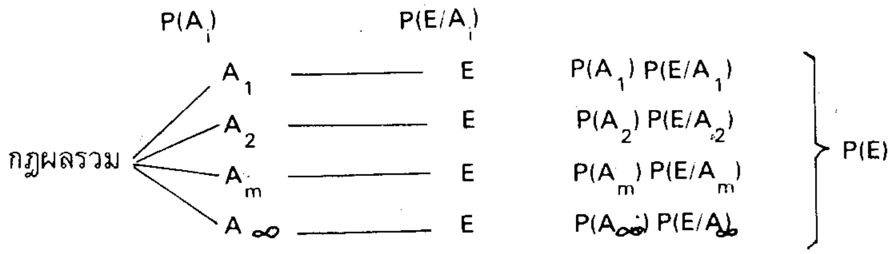
4. กฎผลรวม (Total Probability) ถ้า $A_1, A_2 \ldots$ ประกอบเป็นส่วนแบ่งของกลุ่มผลทดลอง S และ E เป็นเหตุการณ์ใด ๆ ใน S โดยที่ $E = \phi$ แล้ว

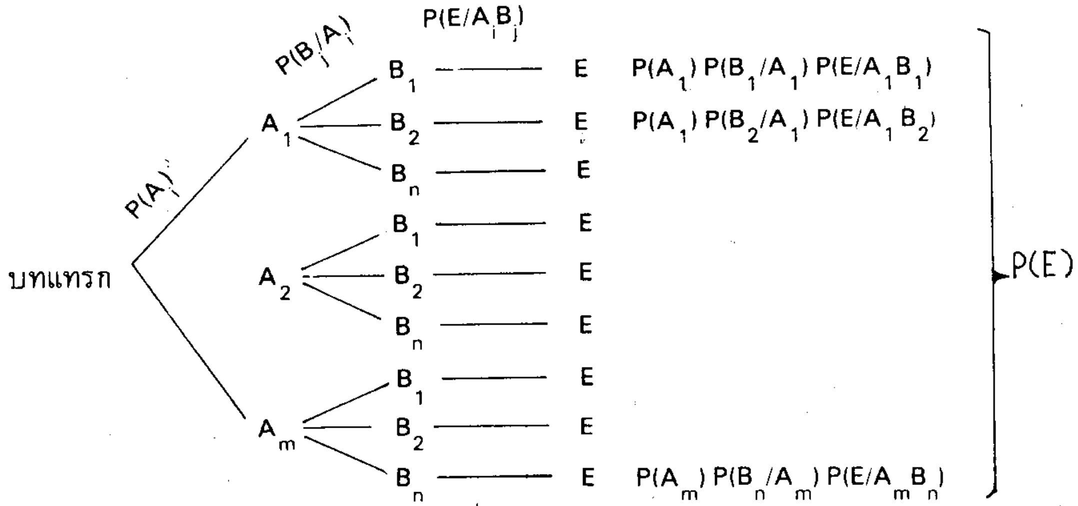
$$P(E) = \sum_i P(E \cap A_i) = \sum_i P(A_i)\ P(E/A_i)$$

บทแทรก ถ้า $A_1, A_2, \ldots, A_m$ และ $B_1, B_2, \ldots, B_n$ ต่างก็ประกอบเป็นส่วนแบ่งของกลุ่มผลทดลอง S และถ้า $E (E \neq \phi)$ เป็นเหตุการณ์ใด ๆ ใน S แล้ว

$$P(E) = \sum_i^m \sum_j^n P(E \cap A_i \cap B_j) = \sum_i \sum_j P(A_i)P(B_j/A_i)P(E/A_iB_j)$$

ในกฎผลรวมและบทแทรก เราสามารถแสดงให้เห็นได้ชัดเจนด้วยแผนภาพ (Tree diagrams) ดังนี้





ตัวอย่าง ในการสำรวจแรงงานพบว่าคนที่อยู่ในวัยทำงานมีระดับการศึกษาคิดเป็นเปอร์เซ็นต์ดังนี้

ประถมศึกษา	50
มัธยมศึกษา	40
อุดมศึกษา	10

และยังพบว่าเปอร์เซ็นต์ของคนในวัยทำงานที่มีระดับการศึกษา จะว่างงานดังนี้

ประถมศึกษาว่างงาน	5
มัธยมศึกษาว่างงาน	2
อุดมศึกษาว่างงาน	0.1

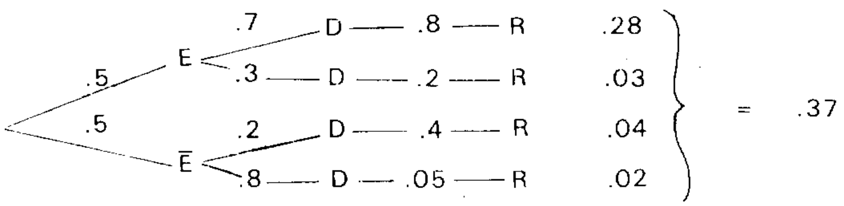
จงหาความน่าจะเป็นของคนว่างงานทั้งหมด

ให้ G_1, G_2 และ G_3 เป็นเหตุการณ์ที่คนในวัยทำงานมีการศึกษาระดับประถม, มัธยม, และอุดมศึกษา ตามลำดับ และ U เป็นเหตุการณ์ที่คนในวัยทำงานจะว่างงาน

$$P(U) = P(G_1)P(U/G_1) + P(G_2)P(U/G_2) + P(G_3)P(U/G_3)$$
$$= .5(.05) + .4(.02) + .1(.001)$$
$$= .025 + .008 + .0001 = .0331$$

ตัวอย่าง ผู้เชี่ยวชาญด้านภัตราคารประมาณว่าในวันสุดสัปดาห์ของต้นเดือน จะมีคนออกไปทานอาหารนอกบ้าน 50% และเป็นที่ทราบกันว่า คนที่ทานอาหารนอกบ้านจะดื่มเหล้า 70% ส่วนผู้ที่ทานอาหารอยู่กับบ้านจะดื่มเหล้า 20% สำหรับผู้ทานอาหารนอกบ้านและดื่มเหล้า โอกาสที่จะไปเที่ยวหาความสำราญต่ออีกเป็น 0.80 และผู้ที่ทานอาหารนอกบ้านแต่ไม่ดื่มเหล้า โอกาสที่จะไปเที่ยวหาความสำราญต่ออีกเป็น 0.20 ส่วนผู้ที่ทานอยู่กับบ้านและดื่มเหล้าโอกาสที่จะไปเที่ยวหาความสำราญเป็น 0.40 และผู้ที่ทานในบ้านและไม่ดื่มเหล้า โอกาสที่จะเที่ยวเป็น 0.05 จงประมาณเปอร์เซ็นต์ที่คนไม่ได้ไปเที่ยวในสุดสัปดาห์ของต้นเดือน

ให้ E แทนเหตุการณ์ที่คนทานอาหารนอกบ้านในสุดสัปดาห์ต้นเดือน
D แทนเหตุการณ์ที่คนดื่มเหล้า
R แทนเหตุการณ์ที่คนไปเที่ยวหาความสำราญ



ดังนั้น ความน่าจะเป็นที่คนจะไม่ไปเที่ยวในวันสุดสัปดาห์

$$= 1 - 0.37 = 0.63$$

5. **ทฤษฎีเบยส์** (Bayes' Theorem) ทฤษฎีเบยส์เป็นกฎที่ใช้กำหนดความน่าจะเป็นเงื่อนไขนั่นเอง แต่เป็นทฤษฎีที่มีความสำคัญมากทฤษฎีหนึ่งโดยเฉพาะในวิชาทฤษฎีตัดสินใจและสถิติแบบเบยส์ (Bayesian Statistics) ทฤษฎีเบยส์กล่าวไว้เป็นขั้น ๆ ดังนี้

(1) ในการทดลองอย่างหนึ่งจะให้เซทของเหตุการณ์ที่เป็นไปได้ที่ไม่ร่วมกัน n เหตุการณ์ นั่นคือ $A_1, A_2, \ldots, A_n$

(2) ความน่าจะเป็นก่อนทดลอง (Prior Probability) ของเหตุการณ์ A_i หรือ $P(A_i)$; i = 1, 2, ..., n นั้นสามารถกำหนดได้โดยอาศัยการกำหนดแบบใดแบบหนึ่ง

(3) นำการทดลองบางอย่างหรือได้ประจักษ์พยานบางอย่างเกี่ยวกับการเกิดขึ้นของเหตุการณ์ E ที่มีเงื่อนไขกับเหตุการณ์ $A_1, A_2, \ldots, A_n$ **นั่นคือเราได้ความน่าจะเป็นเงื่อนไข** $P(E/A)$

(4) เราปรับปรุงความน่าจะเป็นของเหตุการณ์ A_i โดยมีเงื่อนไขว่าเราได้สังเกตการเกิดขึ้นของเหตุการณ์ E แล้ว

ความน่าจะเป็นที่ปรับปรุงนี้เรียกว่า ความน่าจะเป็นหลังทดลอง (Posterior Probability

หรือ $P(A_i/E)$ กำหนดได้จากทฤษฎีเบยส์ ดังนี้

$$P(A_i/E) = \frac{P(A_i)P(E/A_i)}{P(A_1)P(E/A_1) + \cdots + P(A_n)P(E/A_n)}$$

ทฤษฎีเบยส์กล่าวว่า ความน่าจะเป็นหลังทดลองของเหตุการณ์ A หลังจากสังเกตประจักษ์พยาน E นั้น จะเป็นสัดส่วนกับผลคูณของความน่าจะเป็นก่อนทดลอง A_i และความน่าจะเป็นเงื่อนไข (likelihood) ของเหตุการณ์ E เมื่อกำหนด A_i ได้ นั่นคือ

$$P(A_i/E) \propto P(A_i)\,P(B/A_i)$$

ทฤษฎีนี้ Thomas Bayes (1702-1761) นักคณิตศาสตร์และนักปรัชญาชาวอังกฤษได้ประยุกต์จากความน่าจะเป็นเงื่อนไข ซึ่งเขาได้เขียนผลงานไว้ในบทความชื่อ An Essay Toward Solving a Problem in the Doctrine of Chances แต่บทความนี้เพิ่งตีพิมพ์เมื่อ 1763 (หลังมรณกรรม) ใน The Philosophical transactions ของ The Royal Society บทความนี้มีทฤษฎีหรือคำกล่าว (Statement) และวิธีพิสูจน์ของข้อเสนอ (Proposition) ต่อ ๆ มาได้พัฒนาไปอีก จนได้ชื่อว่า Bayes' Theorem (Rule, Formula) ซึ่งในปัจจุบันเป็นที่สนใจกันมากในทฤษฎีตัดสินใจ (Decision Theory)

ข้อเสนอเริ่มแรกเป็นสมการ

$$P(A/E) = \frac{P(E/A)\,P(A)}{P(E)}$$

$$\propto P(A)\,P(E/A)$$

Bayes เป็นคนแรกที่เขียนงานเกี่ยวกับการปรับปรุงความน่าจะเป็น โดยใช้ข้อมูลข่าวสารตัวอย่าง (Sample Information) เหตุการณ์ E นี้ เรียกว่าข้อมูลข่าวสารตัวอย่างทฤษฎีของเบยส์ จึงเป็นการพิจารณาความน่าจะเป็นของเหตุการณ์ที่กำหนดให้ (A_i) ด้วยการใช้ข้อมูลข่าวสารตัวอย่างแบบหนึ่ง

สำหรับเหตุการณ์ $A_1, A_2, \ldots\ldots, A_n$ นั้น ได้ชื่อว่าสมมติฐาน หรือเหตุ (Hypotheses or Causes) และความน่าจะเป็นที่กำหนดให้แก่สมมติฐานนี้ เรียกว่าความน่าจะเป็นก่อนทดลอง (Prior Probabilities) นั่นคือ เหตุการณ์เหล่านี้ได้รับการกำหนดความน่าจะเป็นก่อนที่ข้อมูลข่าวสารใด ๆ จะได้รับจากการทดลอง ความน่าจะเป็นนี้ จะกำหนดโดยยึดถือข้อมูลข่าวสารเชิงปรนัย (Objective Information) หรือประสบการณ์ส่วนตัว (Personal Experience) ที่บอกถึงดีกรีของความเชื่อในสมมติฐานนั้น ซึ่งเป็นข้อมูลข่าวสารเชิงจิตวิสัย (Subjective Information) นั่นเอง

เมื่อได้ทำการทดลองหรือหาข้อมูลข่าวสารตัวอย่างใหม่ ซึ่งจะได้ว่าเหตุการณ์ E เกิดขึ้นแล้วเราก็พิจารณาว่า ความน่าจะเป็นที่กำหนดให้แก่สมมติฐานนั้น ได้เปลี่ยนแปลงไปอย่างไรอันเนื่องมาจากความจริงที่ว่าเหตุการณ์ E ได้เกิดขึ้นแล้ว

จากข้อมูลข่าวสารที่ได้มาใหม่ เราจึงทำการปรับปรุงความน่าจะเป็นก่อนการทดลอง

แล้ว เราจะได้รับความน่าจะเป็นหลังทดลอง (Posterior Probabilities) นั่นคือ เราวางแผนที่จะปรับปรุงความน่าจะเป็นเดิม $P(A_i)$, $i = 1, 2, \ldots, n$ จากประจักษ์พยานใหม่ (New Evidence or Sample Information) เพื่อให้ได้มาซึ่งความน่าจะเป็นที่ปรับปรุงหรือหลังทดลอง (Revised or Posterior Probabilities) หรือ $P(A_i/E)$ ดังนี้

$$P(A_i/E) = P(A_i)P(E/A_i) \Big/ \sum_{j=1}^{n} P(A_j)P(E/A_j) \qquad ; 1 \leq i \leq n$$

ตัวอย่าง พนักงานฝึกหัดขายได้รับการบอกเล่าจากผู้จัดการของเขาว่า 25% ของบุคคลที่เข้ามาในร้านจะซื้อของ และในระหว่างฝึกหัดสองสามเดือนแรก เขาพบว่าหนึ่งในสามของบุคคลที่ซื้อของจะมีบัญชีเงินสดอยู่กับร้าน ส่วนผู้ไม่ได้ซื้อของจะมีบัญชีอยู่กับร้านเพียง 10%

ถ้าบุคคลหนึ่งที่เข้ามาในร้านมีบัญชีเงินสด โอกาสที่บุคคลนี้จะซื้อของเป็นเท่าใด จงหาความน่าจะเป็นที่บุคคลนี้จะไม่ซื้อสิ่งใดเลยจากร้าน

ก. เราตั้งสมมติฐานดังนี้ A_1 : บุคคลที่เข้ามาในร้านจะซื้อของ

A_2 : บุคคลที่เข้ามาในร้านจะไม่ซื้อของ

หตุการณ์ E กำหนดให้เป็นบุคคลที่เข้ามาในร้านมีบัญชีเงินสดอยู่กับร้าน

$\bar{E}$ เป็นบุคคลที่เข้ามาในร้านจะไม่มีบัญชี

ความน่าจะเป็นก่อนการทดลอง ซึ่งกำหนดให้แก่สมมติฐาน เป็นดังนี้

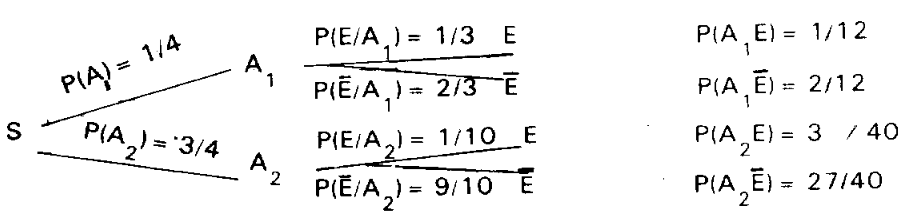
$$P(A_1) = 0.25 = 1/4$$
$$P(A_2) = 0.75 = 3/4$$

ข. ความน่าจะเป็นเงื่อนไขที่กำหนด จากข้อมูลข่าวสารใหม่ ซึ่งได้มาจากการสังเกต หรือทดลองโดยกำหนดว่าสมมติฐานเกิดขึ้นแล้ว ซึ่งเราจะได้ว่า

$$P(E/A_1) = 1/3 \text{ และ } P(E/A_2) = 1/10$$

กลุ่มผลทดลองสามารถแสดงได้ด้วยแผนภาพ ดังนี้

ความน่าจะเป็นก่อนทดลอง **ความน่าจะเป็นเงื่อนไข** ความน่าจะเป็นร่วม



ค. ความน่าจะเป็นร่วมเหล่านี้คำนวณมาได้ดังนี้.

$$P(A_1E) = P(A_1)P(E/A_1) = (1/4)(1/3) = 1/12 = 10/120$$
$$P(A_1\bar{E}) = P(A_1)P(\bar{E}/A_1) = (1/4)(2/3) = 2/12 = 20/120$$

$P(A_2E) = P(A_2)P(E/A_2) = (3/4)(1/10) = 3/40 = 9/120$

$P(A_2\bar{E}) = P(A_2)P(\bar{E}/A_2) = (3/4)(9/10) = 27/40 = 81/120$

ง. ความน่าจะเป็นหลักการทดลองหาได้จากทฤษฎีเบยส์ ดังนี้

$$P(A_1/E) = \frac{P(A_1)P(E/E_1)}{P(A_1)P(E/A_1) + P(A_2)P(E/A_2)}$$

$$= \frac{(1/4)(1/3)}{(1/4)(1/3) + (3/4)(1/10)} = 10/19$$

$P(A_2/E) = 1 - p(A_1/E) = 1 - 10/19 = 9/19$

จาก ก. ถึง ง. เราสรุปรวมได้ดังตารางต่อไปนี้

เหตุการณ์	ความน่าจะเป็นก่อนทดลอง	ความน่าจะเป็นเงื่อนไข	ความน่าจะเป็นร่วม	ความน่าจะเป็นหลังทดลอง
A_1	$P(A_1) = 1/4$	$P(E/A_1) = 1/3$	$P(EA_1) = 1/12$	$P(A_1/E) = 10/19$
A_2	$P(A_2) = 3/4$	$P(E/A_2) = 1/10$	$P(EA_2) = 3/40$	$P(A_2/E) = 9/19$

$P(E) = P(A_1)P(E/A_1) + P(A_2)\ P(E/A_2) = 19/120$

ความน่าจะเป็นเดิมที่กำหนดให้แก่เหตุการณ์ A_1 และ A_2 นั้นอาจจะกำหนดโดยบันทึกที่ร้านทำไว้ หรือความเข้าใจเกี่ยวกับสภาวการณ์ของผู้จัดการเอง หรือได้ทั้งสองอย่างประกอบกัน เมื่อรวมข้อมูลข่าวสารเพิ่มเติมที่ได้จากการสังเกตว่า บุคคลที่เข้ามาในร้านจะมีบัญชีเงินสด จะทำให้ความน่าจะเป็นที่บุคคลนั้นจะซื้อของหรือ $P(A_1/E)$ น่าเชื่อถือมากขึ้น ความน่าจะเป็นเหล่านี้คือความน่าจะเป็นหลังการทดลอง

2.11 ความเป็นอิสระของเหตุการณ์ (Statistical Independence of Events)

บ่อยครั้งเราทราบว่าเหตุการณ์หนึ่งเกิดขึ้นแล้ว แต่ก็ไม่มีผลทำให้ความน่าจะเป็นของอีกเหตุการณ์หนึ่งเปลี่ยนไป นั่นคือ $P(E/E_1) = P(E)$ หรือ $P(E_1/E)$ เราจะพูดว่าสองเหตุการณ์นี้เป็นอิสระต่อกัน ซึ่งเราจะให้นิยามไว้ดังนี้

นิยาม สองเหตุการณ์ใด ๆ จะเรียกว่าเป็นอิสระต่อกัน ถ้าการเกิดขึ้นของเหตุการณ์หนึ่งจะไม่มีผลกระทบต่อการเกิดขึ้นของเหตุการณ์หนึ่ง

ทฤษฎี A และ B เป็นเหตุการณ์ในกลุ่มผลทดลอง S จะเป็นอิสระกันก็ต่อเมื่อ

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

บทแทรก ถ้า $E_1, E_2, \ldots\ldots, E_n$ เป็นเหตุการณ์ใน S ซึ่งเป็นอิสระกันแล้ว

$$P(E_1 \cap E_2 \cap \ldots\ldots \cap E_n) = P(E_1)P(E_2)\ldots\ldots P(E_n)$$

บทกลับของบทแทรกนี้ไม่จริงเสมอไป

ตัวอย่าง เครื่องมือไฟฟ้าอย่างหนึ่งมีส่วนประกอบที่สำคัญอยู่ 2 ส่วน แต่ละส่วนทำงานเป็นอิสระกัน เครื่องมือจะทำงานได้ ถ้าส่วนประกอบอันหนึ่งอันใดทำงานได้ ถ้าความน่าจะเป็นที่ส่วนประกอบ

อันหนึ่งอันใด จะทำงานไม่ได้เป็น 0.01

จงหาความน่าจะเป็นที่เครื่องมือไฟฟ้าจะทำงานไม่ได้

ให้ E_1 และ E_2 เป็นส่วนประกอบตัวแรกและตัวที่สองที่ทำงานไม่ได้

$$P(E_1) = P(E_2) = 0.01$$

ดังนั้น เครื่องมือไฟฟ้าจะทำงานไม่ได้มีความน่าจะเป็น $P(E_1 \cap E_2)$

$$P(E_1 \cap E_2) = P(E_1)P(E_2) = (.01)(.01)$$

$$= .0001$$

ข้อระวัง (1) เหตุการณ์ที่เป็นอิสระกัน กับเหตุการณ์ที่ไม่มีผลร่วมกันนั้น เป็นคนละเรื่องกัน อย่าเอามาปนกัน ความเป็นอิสระ (Independence) เป็นแนวความคิดที่เกี่ยวข้องกับการที่จะกำหนดความน่าจะเป็น แต่ความไม่มีผลร่วมกัน (Disjointedness) ไม่ได้นิยามในเทอมของความน่าจะเป็น แต่นิยามในเทอมของเหตุการณ์

แบบภาพเวนน์ (Venn Diagram) สามารถแสดงให้เห็นถึงคุณสมบัติของความไม่มีผลร่วมกันของเหตุการณ์ได้ แต่เราไม่สามารถแสดงความเป็นอิสระด้วยแผนภาพเวนน์ได้