

ความน่าจะเป็น กับการประยุกต์ใช้ ในการอนุมานเชิงสถิติ

Probability with Application
to Statistical Inference



มานัดน์ คำทอง

ภาควิชาสถิติ คณะวิทยาศาสตร์
มหาวิทยาลัยเชียงใหม่

คำนำ

หนังสือความน่าจะเป็นกับการประยุกต์ใช้ในการอนุมานเชิงสถิติ ผู้เขียนจัดทำขึ้นเพื่อให้ผู้ที่มีความสนใจศึกษาเกี่ยวกับการอนุมานเชิงสถิติ ได้เข้าใจเกี่ยวกับการนำความน่าจะเป็นมาประยุกต์ใช้ในการอนุมานเชิงสถิติ ทำไม่ต้องมีความรู้ความเข้าใจเกี่ยวกับความน่าจะเป็นก่อนที่จะได้ศึกษาการอนุมานเชิงสถิติ เนื้อหาหนังสือประกอบด้วย ความน่าจะเป็นและสถิติ ตัวแปรสุ่ม ฟังก์ชันความน่าจะเป็นและสมบัติที่สำคัญ ฟังก์ชันความน่าจะเป็นร่วมและฟังก์ชันความน่าจะเป็นอย่างมีเงื่อนไข การแจกแจงความน่าจะเป็นของตัวแปรสุ่มไม่ต่อเนื่องและต่อเนื่อง การแจกแจงความน่าจะเป็นของฟังก์ชันตัวแปรสุ่ม ทฤษฎีบทขีดจำกัดเบื้องต้น และองค์ประกอบของการอนุมานเชิงสถิติ โดยที่หนังสือเล่มนี้จะมุ่งเน้นเกี่ยวกับการนำแนวความคิดของการเกิดทฤษฎีการแจกแจงทางสถิติ พร้อมทั้งการประยุกต์ใช้ในการแก้ไขปัญหาการอนุมานเชิงสถิติขั้นสูงต่อไป

ผู้เขียนหวังว่าหนังสือเล่มนี้จะเป็นประโยชน์แก่นักศึกษาและผู้ที่มีความสนใจศึกษาในบางเรื่องที่สนใจศึกษาการประยุกต์ใช้ความน่าจะเป็นต่อการอนุมานเชิงสถิติ ทั้งนี้ผู้เขียนได้เล็งเห็นถึงความสำคัญในการใช้โปรแกรมสำเร็จรูปเพื่อช่วยในการวิเคราะห์ปัญหาด้านสถิติ ที่สามารถเข้าถึงได้ง่ายและสามารถพัฒนาการเขียนโปรแกรมเพื่อตอบสนองการวิเคราะห์ข้อมูลได้

ผู้เขียนขอกราบขอบพระคุณพ่อ-แม่ พี่ๆ น้อง รวมทั้งภรรยาและลูกชายที่สนับสนุน ให้กำลังใจและส่งเสริมการผลิตหนังสือของผู้เขียนมาโดยตลอด กราบขอบพระคุณ ครู-อาจารย์ทุกท่านที่ได้ให้ความรู้แก่ผู้เขียนตลอดมา กราบขอบพระคุณหลักธรรมคำสั่งสอนขององค์สมเด็จพระสัมมาสัมพุทธเจ้าศาสดาของพระพุทธศาสนาที่ผู้เขียนใช้เป็นเครื่องยึดเหนี่ยวจิตใจและนำมาปฏิบัติในชีวิตตลอดเรื่อยมา และขอขอบคุณภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ ที่ให้การสนับสนุนและเอื้อประโยชน์ให้ผู้เขียนได้ผลิตหนังสือเล่มนี้

ผู้เขียนต้องขอบคุณผู้ประเมินหนังสือเล่มนี้ที่ได้ให้ข้อเสนอแนะที่เป็นประโยชน์ต่อการแก้ไขหนังสือเล่มนี้ให้สมบูรณ์มากยิ่งขึ้น อีกทั้งยังขอขอบคุณสำนักพิมพ์มหาวิทยาลัยเชียงใหม่ ที่ได้ให้ความกรุณาในการจัดพิมพ์หนังสือเล่มนี้

หากหนังสือเล่มนี้ มีข้อบกพร่องหรือความผิดพลาดประการใด โปรดแจ้งกลับมาที่ผู้เขียนที่ e-mail: manad.k@cmu.ac.th ผู้เขียนขออภัยเป็นอย่างสูงในข้อบกพร่อง/ความผิดพลาดนั้น และจะได้นำมาปรับแก้ไขในโอกาสต่อไป

มานัดฤทธิ์ คำกอง

มกราคม 2560

สารบัญ

คำนำ	V
คำนิยม	III
สารบัญ	VII
สารบัญตาราง	X
สารบัญภาพ	XI
บทที่ 1 ความน่าจะเป็นและสถิติ	1
1.1 บทนำ	2
1.2 การประยุกต์ใช้ความน่าจะเป็นในการอนุมานเชิงสถิติ	8
1.3 แนวคิดพื้นฐานความน่าจะเป็นและสมบัติความน่าจะเป็น	14
แบบฝึกหัด	23
บรรณานุกรม	26
บทที่ 2 ตัวแปรสุ่ม ฟังก์ชันความน่าจะเป็นและสมบัติที่สำคัญ	27
2.1 บทนำ	28
2.2 ตัวแปรสุ่มและฟังก์ชันความน่าจะเป็น	29
2.3 การคาดหมายและความแปรปรวน	36
2.4 โมเมนต์และฟังก์ชันก่อกำเนิดโมเมนต์	45
แบบฝึกหัด	53
บรรณานุกรม	58
บทที่ 3 ฟังก์ชันความน่าจะเป็นร่วมและฟังก์ชันความน่าจะเป็น	
 อย่างมีเงื่อนไข	59
3.1 บทนำ	60
3.2 ฟังก์ชันความน่าจะเป็นร่วม	61
3.3 ฟังก์ชันความน่าจะเป็นอย่างมีเงื่อนไข	76
3.4 การคาดหมายอย่างมีเงื่อนไขและความแปรปรวนอย่างมีเงื่อนไข	83
แบบฝึกหัด	94
บรรณานุกรม	98

สารบัญ

บทที่ 4	การแจกแจงความน่าจะเป็นของตัวแปรสุ่มไม่ต่อเนื่อง	99
4.1	บทนำ	100
4.2	การแจกแจงเอกรูปไม่ต่อเนื่อง	101
4.3	การแจกแจงแบร์นูลลีและการแจกแจงทวินาม	103
4.4	การแจกแจงเรขาคณิตและการแจกแจงทวินามเชิงลบ	111
4.5	การแจกแจงไฮเพอร์จีโอเมตริก	122
4.6	การแจกแจงปัวซอง	126
4.7	ความสัมพันธ์ระหว่างการแจกแจงไม่ต่อเนื่อง	131
	แบบฝึกหัด	137
	บรรณานุกรม	142
บทที่ 5	การแจกแจงความน่าจะเป็นของตัวแปรสุ่มต่อเนื่อง	143
5.1	บทนำ	144
5.2	การแจกแจงเอกรูป	145
5.3	การแจกแจงแบบเลขชี้กำลังและการแจกแจงแกมมา	149
5.4	การแจกแจงปิตา	163
5.5	การแจกแจงปกติ	167
	แบบฝึกหัด	178
	บรรณานุกรม	182
บทที่ 6	การแจกแจงความน่าจะเป็นของฟังก์ชันตัวแปรสุ่ม	183
6.1	บทนำ	184
6.2	เทคนิคฟังก์ชันความน่าจะเป็นสะสม	185
6.3	เทคนิคการเปลี่ยนตัวแปร	193
6.4	เทคนิคฟังก์ชันก่อกำเนิดโมเมนต์	207
6.5	การแจกแจงของฟังก์ชันตัวแปรสุ่มปกติ	210
	แบบฝึกหัด	225
	บรรณานุกรม	230

สารบัญ

บทที่ 7 ทฤษฎีบทขีดจำกัดเบื้องต้น	231
7.1 บทนำ	232
7.2 ชนิดการลู่เข้า	233
7.3 กฎจำนวนมาก	240
7.4 ทฤษฎีบทขีดจำกัดส่วนกลาง	245
แบบฝึกหัด	258
บรรณานุกรม	262
บทที่ 8 องค์ประกอบของการอนุมานเชิงสถิติ	263
8.1 บทนำ	264
8.2 แนวความคิดเบื้องต้นเกี่ยวกับตัวอย่างสุ่ม	265
8.3 การประมาณค่าแบบจุดและสมบัติที่สำคัญ	272
8.4 การทดสอบสมมุติฐาน	278
8.5 การประมาณค่าแบบช่วงและสมบัติที่สำคัญ	289
แบบฝึกหัด	297
บรรณานุกรม	300
ภาคผนวก	301
ภาคผนวก 1 ทบทวนความรู้เชิงคณิตศาสตร์	302
ภาคผนวก 2 โปรแกรมอาร์เบื้องต้น	305
ภาคผนวก 3 การสร้างตารางความน่าจะเป็นของการแจกแจงทวินาม	310
ภาคผนวก 4 การสร้างตารางความน่าจะเป็นของการแจกแจงปัวซอง	311
ภาคผนวก 5 การสร้างตารางการแจกแจงสะสมปกติมาตรฐาน	312
ภาคผนวก 6 การสร้างตารางการแจกแจงสะสมที่	314
แนวทางอธิบายแบบฝึกหัด	316
ดัชนี	341

สารบัญตาราง

4.1	ค่าคุณลักษณะที่สำคัญของการแจกแจงปัวซอง และการแจกแจงไม่ต่อเนื่องที่เกี่ยวข้อง	134
8.1	ความสัมพันธ์ระหว่างความผิดพลาดแบบที่ 1 และ 2	280

สารบัญภาพ

1.1	ขอบข่ายของการศึกษาสถิติศาสตร์	3
1.2	ชนิดของความน่าจะเป็น	9
2.1	ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X ที่สนใจศึกษา	32
2.2	ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม Y ที่สนใจศึกษา	34
2.3	ลักษณะโค้งความถี่ของข้อมูลที่วัดด้วยค่าสัมประสิทธิ์ความเบ้	46
2.4	ลักษณะโค้งความถี่ของข้อมูลที่วัดด้วยค่าสัมประสิทธิ์ความโด่ง	47
3.1	แนวความคิดของการใช้ทฤษฎีบทของเบส์	79
4.1	ฟังก์ชันความน่าจะเป็นของการแจกแจงเอกรูปไม่ต่อเนื่อง เมื่อ $N = 4\ 6$ และ 10	102
4.2	ฟังก์ชันความน่าจะเป็นของการแจกแจงทวินาม เมื่อ $n = 5\ 15\ 30$ และค่า $p = 0.1\ 0.5\ 0.9$	106
4.3	ฟังก์ชันความน่าจะเป็นของการแจกแจงเรขาคณิต เมื่อ $p = 0.2\ 0.5$ และ 0.8	113
4.4	ฟังก์ชันความน่าจะเป็นของการแจกแจงทวินามเชิงลบ เมื่อ $r = 10\ 20$ และ $p = 0.15\ 0.5\ 0.85$	121
4.5	ฟังก์ชันความน่าจะเป็นของการแจกแจงไฮเพอร์จีโอเมตริก $M=10\ N\ M=10\ 20$ และ $n = 3\ 5\ 7$	125
4.6	ฟังก์ชันความน่าจะเป็นของการแจกแจงปัวซอง เมื่อ $\lambda = 3\ 10$ และ 15	128
4.7	ความสัมพันธ์ระหว่างการแจกแจงปัวซอง และการแจกแจงไม่ต่อเนื่องที่เกี่ยวข้อง	136
5.1	ฟังก์ชันความน่าจะเป็นของการแจกแจงเอกรูปต่อเนื่อง	147
5.2	ฟังก์ชันความน่าจะเป็นของการแจกแจงแบบเลขชี้กำลัง ด้วยพารามิเตอร์ $\lambda = 1\ 2$ และ 3	152
5.3	ฟังก์ชันความน่าจะเป็นของการแจกแจงแกมมา ด้วยพารามิเตอร์ $\alpha = 1\ 2$ และ 3	158

สารบัญภาพ

5.4	ฟังก์ชันความน่าจะเป็นของการแจกแจงไคกำลังสอง ด้วยองศาเสรี $v = 15$ และ 30	162
5.5	ฟังก์ชันความน่าจะเป็นของการแจกแจงบีตา ด้วยพารามิเตอร์ a และ b	166
5.6	ฟังก์ชันความน่าจะเป็นของการแจกแจงปรกติด้วยค่าเฉลี่ย 1 และส่วนเบี่ยงเบนมาตรฐาน 1 2 และ 3	171
5.7	พื้นที่ใต้ส่วนโค้งการแจกแจงปรกติที่ห่างจากค่าเฉลี่ย เป็นระยะ σ 2σ และ 3σ	172
5.8	การแจกแจงปรกติสองตัวแปรด้วยค่าเฉลี่ย $\mu_1 = \mu_2 = 0$ ความแปรปรวน $\sigma_{11} = \sigma_{22} = 10$ และค่าสัมประสิทธิ์สหสัมพันธ์ $\rho = 0$ และ $\rho = 0.75$	177
6.1	การสร้างการแจกแจงความน่าจะเป็นของฟังก์ชันตัวแปรสุ่ม $Y = g(X)$	185
6.2	ฟังก์ชันการแจกแจงสะสมของ $Y = g(X)$ กรณี $g(\cdot)$ เป็นฟังก์ชันต่อเนื่องเพิ่มโดยแท้จริง	196
6.3	ฟังก์ชันการแจกแจงสะสมของ $Y = g(X)$ กรณี $g(\cdot)$ เป็นฟังก์ชันต่อเนื่องลดโดยแท้จริง	196
6.4	ฟังก์ชันความน่าจะเป็นของ $f(x) = \frac{x^2}{2}$, $-1 < x < 3$	200
6.5	ฟังก์ชันความน่าจะเป็นของการแจกแจงที่ และการแจกแจงปรกติมาตรฐาน	220
6.6	ฟังก์ชันความน่าจะเป็นของการแจกแจงเอฟ	224
8.1	การเลือกสถิติทดสอบค่าเฉลี่ย 1 ประชากรของการแจกแจงปรกติ	283
8.2	อาณาเขตวิกฤตของแต่ละรูปแบบการทดสอบสมมุติฐาน	284

สถิติศาสตร์ เป็นวิทยาศาสตร์ของการเรียนรู้จากข้อมูล การอนุมานเชิงสถิติอาศัยสารสนเทศที่ได้จากตัวอย่างสุ่ม เพื่อสรุปเกี่ยวกับค่าพารามิเตอร์ที่ไม่ทราบค่าที่ผู้สนใจศึกษามีข้อสมมุติเกี่ยวกับฟังก์ชัน การแจกแจงความน่าจะเป็นของ สิ่งที่สนใจศึกษาไว้ หนังสือเล่มนี้มุ่งเน้น การนำความน่าจะเป็น และสมบัติที่สำคัญ ที่จำเป็นต่อการอนุมานสถิติขั้นสูงต่อไป ไม่ว่าจะเป็นการประมาณค่าพารามิเตอร์แบบจุดและแบบ ช่วง การตั้งสมมุติฐานและการทดสอบสมมุติฐาน พร้อมทั้งการ เสนอแนวทางในการประยุกต์ใช้ความน่าจะเป็น ในการอนุมาน เชิงสถิติของผู้ใช้ชนเป็นตัวอย่างประกอบ และการใช้โปรแกรม อาร์ช่วยอธิบายนิยาม สามารถใช้หนังสือเล่มนี้ เพื่อประกอบ การศึกษาและการค้นคว้าเพิ่มเติมได้



CHIANG MAI
UNIVERSITY PRESS

ISBN: 978-616-398-236-0



9 786163 982360

ราคา 400 บาท

ความน่าจะเป็นและสถิติ

PROBABILITY AND STATISTICS

► PROBABILITY WITH APPLICATION TO STATISTICAL INFERENCE

1.1 บทนำ (Introduction)

ความน่าจะเป็นและสถิติคืออะไร? ทำไมต้องศึกษาความน่าจะเป็น?

ในชีวิตประจำวันเราได้รับข้อมูลและข่าวสารต่างๆ อย่างมากมาย ตัวอย่างเช่น การรับฟังรายงานพยากรณ์อากาศประจำวันจากกรมอุตุนิยมวิทยาเพื่อให้ได้รับรู้ข้อมูลสภาพภูมิอากาศเพื่อช่วยในการวางแผนด้านการเดินทาง การรับฟังราคาซื้อขายทองคำต่อวันจากสมาคมค้าทองแห่งประเทศไทยเพื่อให้เห็นแนวโน้มของราคาทองคำ เป็นต้น ดังนั้นเราจำเป็นต้องมีการวิเคราะห์ข้อมูลที่ได้รับข่าวสารมาเพื่อช่วยในการตัดสินใจในหลายๆ ด้านไม่ว่าจะช่วยในการดำรงชีวิตประจำวัน การตัดสินใจที่ประกอบการลงทุนต่างๆ ซึ่งต้องอาศัยการศึกษาศาสตร์ที่เกี่ยวข้องกับสถิติ โดยที่สถิติศาสตร์ (statistics) เป็นวิทยาศาสตร์ที่เกี่ยวข้องกับการวางแผนการศึกษา การเก็บรวบรวมข้อมูล (หรือการสร้างแบบจำลอง) และการวิเคราะห์ข้อมูลเพื่อใช้ประโยชน์ในการตัดสินใจและการค้นพบองค์ความรู้ทางวิทยาศาสตร์ภายใต้ข้อจำกัดด้านข้อมูลและตัวแปรที่ศึกษา นั่นคือ **สถิติศาสตร์ เป็นวิทยาศาสตร์ของการเรียนรู้จากข้อมูล** (Ott และ Longnecker, 2010: 2) ดังนั้น การอนุมานเชิงสถิติอาศัยสารสนเทศที่ได้จากตัวอย่างสุ่ม $X_1, X_2, \dots, X_n$ เพื่อสรุปเกี่ยวกับค่าพารามิเตอร์ที่ไม่ทราบค่า คือคาดการณ์ว่าตัวอย่างสุ่มได้เก็บข้อมูลสารสนเทศที่สำคัญเกี่ยวกับพารามิเตอร์ที่สนใจศึกษาไว้หมดแล้ว ซึ่งภายใต้การสุ่มตัวอย่างมาศึกษาขนาด n ที่ถูกสุ่มมาจากประชากรที่สนใจศึกษาที่มีฟังก์ชันการแจกแจงความน่าจะเป็น $f_X(X|\theta)$ โดยทั่วไปจะไม่ทราบรูปแบบฟังก์ชันความน่าจะเป็นที่แท้จริง แต่จะมีการสมมุติว่าประชากรที่สนใจศึกษามีลักษณะรูปแบบฟังก์ชันความน่าจะเป็นลักษณะอย่างใดอย่างหนึ่งที่มีพารามิเตอร์ที่สนใจศึกษาไม่ทราบค่า θ ซึ่งจะได้ศึกษาอย่างละเอียดในบทถัดไป หากในกระบวนการอนุมานเชิงสถิติที่จำเป็นต้องระบุการแจกแจงของตัวอย่างสุ่มที่มีค่าพารามิเตอร์ที่สนใจศึกษากำกับอยู่จะเรียกการอนุมานนี้ว่า สถิติศาสตร์อิงพารามิเตอร์ (parametric

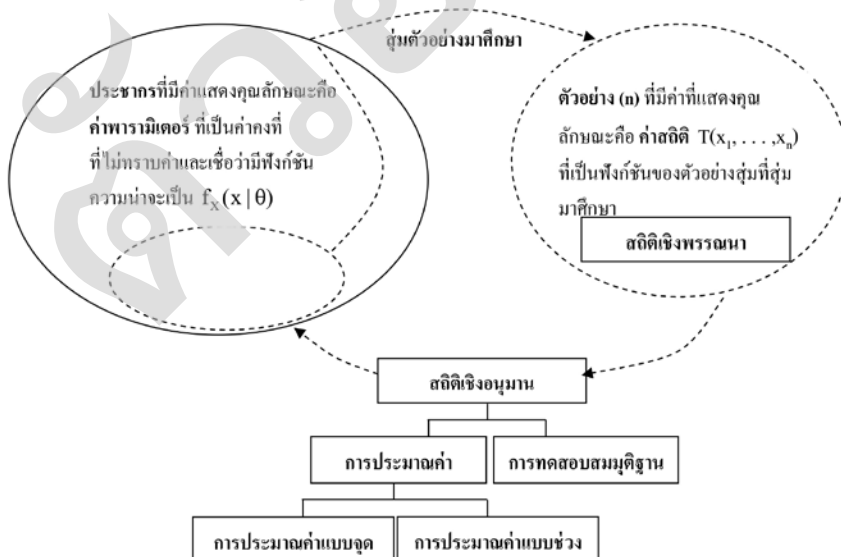
Statistics) และหากการอนุมานนั้นไม่จำเป็นต้องระบุการแจกแจงของตัวอย่างสุ่ม จะเรียกการอนุมานนั้นว่า สถิติศาสตร์ไม่อิงพารามิเตอร์ (non-parametric Statistics) กรณีศึกษาของผู้จัดการฝ่ายขายเครื่องโทรศัพท์มือถือยี่ห้อหนึ่ง ต้องการศึกษากำหนดจำนวนเครื่องโทรศัพท์มือถือยี่ห้อนี้ที่จำหน่ายให้ผู้บริโภคจะเสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี จะมีมากน้อยเพียงใด ดังนั้นในการศึกษานี้จะสนใจประมาณค่าพารามิเตอร์คือสัดส่วน (proportion, p) ของเครื่องโทรศัพท์มือถือของบริษัทจะเสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี ในการศึกษานี้ผู้จัดการฝ่ายขายอาจสุ่มข้อมูลที่ขายเครื่องโทรศัพท์มือถือยี่ห้อนี้จำนวน n มาศึกษา นั่นคือ

$$x = \begin{cases} 0 & \text{แทนโทรศัพท์มือถือที่ไม่เสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี ด้วยความน่าจะเป็น } 1 - p \\ 1 & \text{แทนโทรศัพท์มือถือที่เสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี ด้วยความน่าจะเป็น } p \end{cases}$$

ตัวอย่างสุ่มที่สนใจศึกษาจะไม่ทราบการแจกแจงที่แท้จริงของจำนวนเครื่องโทรศัพท์มือถือที่เสียแต่จากลักษณะของการศึกษานี้ ถ้าผู้จัดการฝ่ายขายจะสมมติให้ตัวแปรสุ่ม X เป็นจำนวนเครื่องโทรศัพท์มือถือที่เสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี จะมีลักษณะการแจกแจงแบร์นูลลี (Bernoulli distribution) ที่มีฟังก์ชันความน่าจะเป็น ดังนี้

$$P(X = x | p) = p^x (1 - p)^{1-x}, \quad x = 0, 1; \quad 0 \leq p \leq 1$$

และศูนย์ที่อื่นๆ โดยที่ไม่ทราบค่าพารามิเตอร์ p ที่สนใจศึกษาแต่ทราบปริภูมิพารามิเตอร์ Θ คือ $p \in \Theta = \{p : 0 \leq p \leq 1\}$ จะเรียกการอนุมานนี้ว่า สถิติศาสตร์อิงพารามิเตอร์ ดังภาพที่ 1.1



ภาพที่ 1.1 ขอบข่ายของการศึกษาสถิติศาสตร์

ก่อนที่จะศึกษาขอบข่ายของศาสตร์สถิติจะอธิบายคำนิยามศัพท์ที่เกี่ยวข้องกับสถิติ ดังนี้

1) ประชากร (population) หมายถึง ชุดขององค์ประกอบหรือขอบเขตทั้งหมดที่สนใจศึกษา อาจเป็นคน สัตว์ หรือสิ่งของก็ได้ สามารถแบ่งเป็น 2 ประเภท คือ

1.1) ประชากรที่ประกอบด้วยชุดขององค์ประกอบหรือขอบเขตทั้งหมดที่สนใจศึกษามีจำนวนนับที่แน่นอน เรียกว่า ประชากรอันตะ (finite population) เช่น จำนวนพนักงานที่ปฏิบัติงานในโรงงานผลิตชิ้นส่วนประกอบยานยนต์ชนิดหนึ่งในช่วงเดือนธันวาคม พ.ศ. 2559 จำนวนค่าที่ปรากฏในหนังสือความน่าจะเป็นกับการประยุกต์ใช้ในการอนุมานเชิงสถิติ เป็นต้น

1.2) ประชากรที่ประกอบด้วยชุดขององค์ประกอบหรือขอบเขตทั้งหมดที่สนใจศึกษามีจำนวนนับที่ไม่แน่นอน เรียกว่า ประชากรอนันต์ (infinite population) เช่น จำนวนปลาที่อาศัยอยู่ในทะเลทั่วโลก จำนวนนกที่อาศัยอยู่ในประเทศไทย เป็นต้น

2) ตัวอย่าง (sample) หมายถึง ชุดขององค์ประกอบหรือขอบเขตบางส่วนของประชากรที่สนใจศึกษา โดยการเลือกมาศึกษาเพื่อเป็นตัวแทนในการหาข้อสรุปเกี่ยวกับประชากรนั้นๆ

3) พารามิเตอร์ (parameter) หมายถึง ค่าคงที่ที่แสดงคุณลักษณะบางประการของประชากร เช่น การทดลองวิธีการใหม่เพื่อเพิ่มผลผลิตข้าวต่อไร่ของเกษตรกรอำเภอท่านหนึ่ง ที่คาดการณ์ผลผลิตข้าวต่อไร่ด้วยวิธีการใหม่มีการแจกแจงปกติที่มีค่าเฉลี่ยของประชากร (population mean) μ และความแปรปรวนของประชากร (population variance) σ^2 เป็นต้น

4) ค่าสถิติ (statistic) หมายถึง ฟังก์ชันของค่าสังเกตที่วัดจากหน่วยต่างๆ ที่ถูกเลือกขึ้นมาเป็นตัวอย่าง ซึ่งจะไม่มีค่าพารามิเตอร์อื่นๆ ที่ไม่ทราบค่าปนอยู่ ค่าสถิติที่ได้จะใช้เป็นตัวประมาณค่าของพารามิเตอร์ เช่น จากการทดลองปลูกข้าวด้วยวิธีการใหม่ของเกษตรกรอำเภอท่านหนึ่งในพื้นที่ต่างๆ อย่างละ 1 ไร่ จำนวนทั้งหมด 30 ไร่เป็นตัวอย่างในการศึกษาภายใต้การควบคุมปัจจัยต่างๆ ให้เหมือนกัน เมื่อเสร็จสิ้นการทดลองได้ทำการวัดผลผลิตแต่ละไร่ โดยการหาค่าเฉลี่ยของตัวอย่าง (sample mean) $\bar{X}$ และความแปรปรวนของตัวอย่าง (sample variance) S^2

5) หน่วยตัวอย่าง (sampling unit) หมายถึง หน่วยที่ต้องการศึกษาหรือจัดรายละเอียดข้อเท็จจริงบางประการ อาจจะเป็นหน่วยเล็กๆ เพียงหน่วยเดียว หรือเป็นกลุ่มของหน่วยเล็กๆ หลายๆ หน่วยรวมกันก็ได้

6) หน่วยแจงนับ (enumeration unit) หมายถึง หน่วยหรือสิ่งของที่จะเก็บรวบรวมข้อมูลในบางกรณีมีความหมายคล้ายกับหน่วยตัวอย่าง เช่น การศึกษาค่าใช้จ่ายต่อเดือนของนักศึกษามหาวิทยาลัยเชียงใหม่ ประจำปีการศึกษา 2559 หน่วยตัวอย่าง คือ นักศึกษามหาวิทยาลัยเชียงใหม่ที่มีสภาพการเป็นนักศึกษาแต่ละคนในปีการศึกษา 2559 เนื่องจากการสำรวจสามารถสอบถามค่าใช้จ่ายต่อเดือนของนักศึกษาแต่ละคนได้ ดังนั้น นักศึกษาแต่ละคนเป็นหน่วยแจงนับด้วย แต่ในบางกรณีหน่วยตัวอย่างและหน่วยแจงนับมีความหมายที่แตกต่างกัน เช่น การสำรวจผลผลิตข้าวของ

ครัวเรือนเกษตรกรในจังหวัดเชียงใหม่ ปี พ.ศ. 2559 โดยกำหนดให้ครัวเรือนเกษตรกรเป็นหน่วยตัวอย่าง โดยที่ครัวเรือนไม่สามารถให้ข้อมูลได้ ดังนั้น จำเป็นต้องถามบุคคลในครัวเรือน เช่น หัวหน้าครัวเรือนแต่ละครัวเรือน ดังนั้นหัวหน้าครัวเรือนเป็นหน่วยแขนง

7) กรอบตัวอย่าง (sampling frame) หมายถึง บัญชีรายชื่อของหน่วยตัวอย่างที่ประกอบด้วยรายละเอียดเกี่ยวกับตำแหน่งที่อยู่ของหน่วยตัวอย่างเหล่านั้น

8) แผนแบบการเลือกตัวอย่าง (sample design) หมายถึง วิธีการเลือกหน่วยตัวอย่างต่าง ๆ เพื่อให้ได้ตัวอย่างที่เป็นตัวแทนของประชากรที่สนใจศึกษา

9) สถิติเชิงพรรณนา (descriptive statistics) หมายถึง การใช้ระเบียบวิธีการทางสถิติเพื่อจัดทำข้อสรุปคุณลักษณะของสิ่งที่สนใจศึกษา โดยไม่อาศัยวิธีการอนุมานไปสู่ประชากร เช่น การวัดแนวโน้มเข้าสู่ส่วนกลางและการวัดการกระจายของข้อมูล เป็นต้น

10) สถิติเชิงอนุมาน (inferential statistics) หมายถึง การใช้ระเบียบวิธีการทางสถิติเพื่อจัดทำข้อสรุปของตัวอย่างที่สุ่มมาศึกษาเพื่ออนุมานไปสู่ประชากรโดยการประเมินความน่าเชื่อถือของวิธีการของสถิติดังกล่าว

ตัวอย่างที่ 1.1 นักวิจัยสิ่งแวดล้อมท่านหนึ่งได้รับข้อร้องเรียนจากชาวบ้านเกี่ยวกับโรงงานผลิตสารกำจัดศัตรูพืชแห่งหนึ่งที่ตั้งอยู่ใกล้กับแม่น้ำได้ลักลอบปล่อยน้ำเสียลงสู่แม่น้ำ ทำให้สิ่งมีชีวิตที่อาศัยอยู่ในแม่น้ำตายเป็นจำนวนมากและแม่น้ำเน่าเสียไม่สามารถอุปโภคได้ นักวิจัยท่านนี้จึงได้สุ่มตรวจระบบบำบัดน้ำเสียของโรงงานแห่งนี้ปริมาณ 500 มิลลิลิตร เพื่อตรวจวัดค่า pH, SS, BOD, FOG และ TDS จากน้ำเสียที่ปล่อยจากโรงงานแห่งนี้มาจำนวน 30 วัน เพื่อตรวจสอบค่าวัดเหล่านั้นจะเกินกว่าที่คณะกรรมการควบคุมมลพิษกำหนดหรือไม่ ดังนี้

ค่าวัด	ค่ามาตรฐาน
• pH คือ ค่าความเป็นกรด-ด่าง	• อยู่ระหว่าง 5.5 – 9 pH meter
• BOD (Biochemical Oxygen Demand) คือ ปริมาณของออกซิเจนที่แบคทีเรียใช้ย่อยสลายสารอินทรีย์ใน 5 วัน ที่ 20 องศาเซลเซียส	• ไม่เกิน 20 มก/ล หรือขึ้นอยู่กับประเภทอุตสาหกรรมแต่ไม่เกิน 60
• SS (Suspended Solids) คือค่าสารแขวนลอย	• ไม่เกิน 50 มก/ล
• FOG (Fat, Oil and Grease) คือ ค่าน้ำมันและไขมัน	• ไม่เกิน 5 มก/ล หรือไม่เกิน 15 มก/ล
• TDS (Total Dissolved Solids) คือ ของแข็งที่ละลายได้ทั้งหมด	• ไม่เกิน 3,000 มก/ล หรือไม่เกิน 5,000 มก/ล

ที่มา: กรมควบคุมมลพิษ กระทรวงทรัพยากรธรรมชาติและสิ่งแวดล้อม

หน่วยวัด: มิลลิกรัมต่อลิตร (มก/ล)

ประชากรที่สนใจศึกษา คือ น้ำเสียที่ปล่อยจากโรงงานแห่งนี้ทั้งหมด

ตัวอย่าง คือ ปริมาณน้ำเสียที่ปล่อยจากโรงงานแห่งนี้ที่สุ่มมาตรวจในแต่ละวันปริมาณ 500 มิลลิลิตร จำนวน 30 วัน

หน่วยตัวอย่าง คือ น้ำเสียที่ปล่อยจากโรงงานแห่งนี้ในแต่ละวัน

หน่วยแจกนับ คือ เนื่องจากปริมาณน้ำเสียที่ปล่อยจากโรงงานแห่งนี้สามารถวัดค่าต่าง ๆ ได้ ดังนั้น หน่วยแจกนับจะเหมือนกับหน่วยตัวอย่าง

พารามิเตอร์ คือ ค่าเฉลี่ย (μ) และความแปรปรวนของประชากร (σ^2) ของค่าวัด pH, SS, BOD, FOG และ TDS หากสมมุติให้ปริมาณน้ำเสียที่ปล่อยออกมาจากโรงงานแห่งนี้ทั้งหมดมีการแจกแจงปกติ

ค่าสถิติ คือ ค่าเฉลี่ย ($\bar{X}$) และความแปรปรวนของตัวอย่าง (S^2) ของค่าวัด pH, SS, BOD, FOG และ TDS ในปริมาณน้ำเสียที่ปล่อยออกมาจากโรงงานที่สุ่มมาตรวจสอบ ในแต่ละวันปริมาณ 500 มิลลิลิตร จำนวน 30 วัน

จากตัวอย่างที่ 1.1 จะเห็นได้ว่าการใช้ตัวอย่างสุ่มเพียงบางส่วนมาศึกษาเพื่อกลับไปอ้างอิงถึงประชากรทั้งหมดนั้นซึ่งอาจมีความผิดพลาดที่อาจเกิดขึ้นได้ ดังนั้นจะต้องใช้ความระมัดระวังในการสรุปผล ในการทดลองหลายๆ สาขาวิชาทางวิทยาศาสตร์ ก็จะทำการศึกษาในลักษณะนี้เช่นกัน แต่จะมีการควบคุมปัจจัยหรือสภาพแวดล้อมในการทดลองตามวัตถุประสงค์ของการศึกษา ซึ่งการอนุมานในลักษณะนี้เรียกว่าการอนุมานอุปนัย (inductive inference)

การอนุมานอุปนัยเป็นกระบวนการที่มีความเสี่ยงต่อการสรุปผลการศึกษาย่างยิ่ง เพราะในความเป็นจริงแล้ว การอนุมานอุปนัยอาศัยตรรกะที่มีความไม่แน่นอนช่วยในการอนุมาน ซึ่งความไม่แน่นอนนั้น สามารถที่จะวัดระดับของความไม่แน่นอนได้ หากการทดลองได้รับการดำเนินการตามหลักการที่ถูกต้อง สำหรับบทบาทหนึ่งของสถิติที่มีผลต่อเทคนิคการทำข้อสรุปแบบอุปนัยนั้นคือการวัดระดับของความไม่แน่นอนของการวินิจฉัยดังกล่าวด้วยความน่าจะเป็น (probability) ดังนั้นความไม่แน่นอนจะถูกวัดออกมาในแง่ของความน่าจะเป็น นั่นเป็นเหตุผลที่สำคัญที่ควรมีการศึกษาทฤษฎีของความน่าจะเป็นเพื่อประยุกต์ใช้ในทฤษฎีการอนุมานเชิงสถิติต่อไป

สำหรับการอนุมานอีกแบบหนึ่งที่นิยมใช้ในการพิสูจน์ทฤษฎีบทความเป็นจริงทางคณิตศาสตร์นั้นคือ **การอนุมานนิรนัย (deductive inference)** โดยที่การอนุมานนิรนัยจะอาศัยเหตุปัจจัยของความเป็นจริงเพื่อเป็นข้อสรุปผลจากเหตุที่ได้

ตัวอย่างที่ 1.2 พิจารณาข้อความจริง 2 ประการดังนี้

- 1) **หลักฐานหลัก** : มุมภายในหนึ่งมุมของรูปสามเหลี่ยมมุมฉากแต่ละรูปเท่ากับ 90 องศา
- 2) **หลักฐานรอง** : สามเหลี่ยม B เป็นรูปสามเหลี่ยมมุมฉากถ้ายอมรับหลักฐานทั้งสองอย่างนี้ แล้วจะถูกต้องให้สรุปว่า
- 3) **ข้อสรุป** : หนึ่งในมุมของรูปสามเหลี่ยม B เท่ากับ 90 องศา

จากตัวอย่างที่ 1.2 จะเห็นได้ว่าเป็นการอนุมานนिरनัยที่ข้อสรุปความจริง 3) จะเกิดจากข้อเท็จจริงที่เรายอมรับหลักฐานจาก 1) และ 2) นั่นคือ การอนุมานนिरนัยเป็นสิ่งสำคัญอย่างมากสำหรับการแสวงหาองค์ความรู้ใหม่ในด้านวิทยาศาสตร์ โดยเฉพาะสาขาคณิตศาสตร์การใช้การอนุมานนिरนัยที่จะช่วยในการพิสูจน์ทฤษฎีบทต่างๆ ในขณะที่การอนุมานเชิงประจักษ์ทางวิทยาศาสตร์จะใช้ออกอนุมานอุปนัยเพื่อการค้นหาคำตอบใหม่

ตัวอย่างที่ 1.3 สมมติว่าร้านขายเมล็ดพันธุ์เบญจมาศได้มีการเก็บเมล็ดพันธุ์เบญจมาศ 2 ชนิด คือ ชนิดสีขาวและสีเหลือง เพื่อจำหน่ายให้แก่เกษตรกรผู้เลี้ยงเบญจมาศได้เกิดอุบัติเหตุทำให้เมล็ดเบญจมาศทั้งสองสายพันธุ์ผสมกัน ซึ่งเมล็ดเบญจมาศของสองสายพันธุ์รวมกันเป็น 5 ล้านเมล็ด โดยที่เมล็ดเบญจมาศของสองสายพันธุ์มีลักษณะใกล้เคียงกัน ซึ่งการจะจำแนกออกเป็นสายพันธุ์ได้นั้น จะต้องมีการทดลองปลูกเมล็ดเบญจมาศจนกว่าจะออกดอกจึงจะสามารถจำแนกสายพันธุ์ได้ **จากข้อมูลดังกล่าวหากมีเกษตรกรมาขอซื้อเมล็ดเบญจมาศไปปลูกจากเจ้าของร้านและถามโอกาสที่จะได้เมล็ดเบญจมาศชนิดสีขาวจากเมล็ดทั้งหมด 5 ล้านเมล็ด?**

จากเหตุการณ์นี้จะเห็นได้ว่ามีความเป็นไปได้น้อยมากที่จะทำการทดลองปลูกเมล็ดเบญจมาศทั้งหมด 5 ล้านเมล็ดและรอจนกว่าจะออกดอกเพื่อจำแนกชนิดซึ่งอาจใช้เวลานาน ดังนั้นแนวทางหนึ่งที่จะได้คำตอบคร่าวๆ อาจมีการสุ่มเมล็ดพันธุ์เบญจมาศไปปลูกเพียงบางส่วนแล้วสรุปผลว่าเป็นเบญจมาศชนิดสีขาวอยู่ร้อยละเท่าไร ซึ่งกระบวนการอนุมานนี้เรียกว่าการอนุมานอุปนัย นั่นคือ คำตอบที่ได้จะเป็นคำตอบที่ไม่ได้แน่ชัดว่าเป็นเบญจมาศชนิดสีขาวจริงๆ แต่เป็นเพียงการประมาณค่าภายใต้สถานการณ์ที่ทดลองเท่านั้น ซึ่งมีหลากหลายปัจจัยที่จะส่งผลต่อการสรุปผล เช่น ตัวอย่างสุ่มที่ศึกษามีขนาดที่เหมาะสมและเป็นตัวแทนที่ดีของประชากรที่สนใจศึกษาหรือไม่ แต่การอนุมานอุปนัยก็มีประโยชน์ในหลายๆ ด้าน เช่น การประหยัดเวลาและค่าใช้จ่าย เป็นต้น

ดังนั้น ปัญหาในการอนุมานอุปนัยได้ถูกพัฒนาเข้ามาสู่กระบวนการอนุมานเชิงสถิติ โดยเริ่มต้นจากการตั้งวัตถุประสงค์ของการศึกษาจากประชากรเป้าหมายที่สนใจศึกษาที่แท้จริง ซึ่งอาจทำการศึกษาจากบางส่วนเกี่ยวกับประชากรเป้าหมาย หรือที่เรียกว่า ตัวอย่างสุ่ม เพราะบางครั้งมีข้อกำหนดด้านเวลาและค่าใช้จ่ายในการศึกษาที่ไม่สามารถศึกษาเต็มประชากรเป้าหมายได้และบนพื้นฐานของการตรวจสอบบางอย่างอาจมีข้อจำกัดด้านตัวแปรที่สนใจศึกษา เพื่อหาข้อสรุปเกี่ยวกับประชากรเป้าหมายทั้งหมดได้ ซึ่งในบทนี้จะได้ศึกษาเกี่ยวกับการประยุกต์ใช้ความน่าจะเป็นในการอนุมานเชิงสถิติ แนวความคิดพื้นฐานของความน่าจะเป็นและสมบัติที่สำคัญต่อการอนุมานเชิงสถิติ

1.2 การประยุกต์ใช้ความน่าจะเป็นในการอนุมานเชิงสถิติ

[Applications of Probability in Statistical Inference]

ความน่าจะเป็น (probability) คือ ตัวเลขที่วัดโอกาสการเกิดของเหตุการณ์ ว่ามีโอกาสเกิดขึ้นได้มากน้อยเพียงใด นั่นคือ ถ้าความน่าจะเป็นของเหตุการณ์มีค่าเท่ากับ 0 หมายความว่าเหตุการณ์นั้นไม่เกิดขึ้น หรือความน่าจะเป็นเท่ากับ 0.5 หมายความว่าเหตุการณ์นั้นมีโอกาสเกิดขึ้น 50% หรือความน่าจะเป็นเท่ากับ 1 หมายความว่าเหตุการณ์นั้นมีโอกาสเกิดขึ้นอย่างแน่นอน ตามลำดับ โดยที่ความน่าจะเป็นที่คำนวณได้มาจากเหตุการณ์หรือการทดลองที่ไม่สามารถทำนายหรือบอกได้ล่วงหน้าว่าจะเกิดเหตุการณ์ใดอย่างแน่ชัด เรียกว่า **การทดลองสุ่ม (random experiment)** ดังนั้น ความน่าจะเป็นคือฟังก์ชันค่าจริงจากการเปลี่ยนเซตของผลลัพธ์ที่เป็นไปได้ทั้งหมดของการทดลองสุ่มสู่จำนวนจริงที่มีค่าตั้งแต่ 0 ถึง 1 ภายใต้สัจพจน์บางอย่าง

ปริภูมิตัวอย่าง (sample space) คือ เซต หรือกลุ่มของผลลัพธ์ที่เป็นไปได้ทั้งหมดจากการทดลองสุ่ม แทนด้วย S หรือ Ω

เหตุการณ์ (event) คือ สมาชิกหรือกลุ่มของสมาชิกที่สนใจซึ่งเป็นส่วนหนึ่งในปริภูมิตัวอย่าง

สำหรับการคำนวณความน่าจะเป็นของเหตุการณ์สามารถจำแนกได้เป็น 3 แนวทาง ดังภาพ 1.2 ได้แก่ 1) ความน่าจะเป็นเชิงคลาสสิก (classical probability) 2) ความน่าจะเป็นเชิงประจักษ์ (empirical probability) และ 3) ความน่าจะเป็นเชิงจิตวิสัย (subjective probability) โดยที่ความน่าจะเป็นพื้นฐานของเหตุการณ์ใดๆ จะอยู่ภายใต้สองเงื่อนไขที่สำคัญ ดังนี้

1) ความน่าจะเป็นของแต่ละเหตุการณ์ของการทดลองสุ่มจะมีค่าอยู่ระหว่าง 0 ถึง 1 เช่น ถ้าให้ E_i แทนผลลัพธ์ในเหตุการณ์ที่ i ของการทดลองสุ่ม และ $P(E_i)$ แทนความน่าจะเป็นของเหตุการณ์ E_i สามารถเขียนได้ดังนี้

$$0 \leq P(E_i) \leq 1, \text{ สำหรับทุกๆ เหตุการณ์ } i, i = 1, 2, 3 \dots, n$$

2) ผลรวมของความน่าจะเป็นสำหรับทุกผลลัพธ์ในเหตุการณ์ที่เป็นไปได้ทั้งหมดของการทดลองสุ่มเท่ากับ 1; $P(E_1) + P(E_2) + \dots + P(E_m) = 1$



ภาพที่ 1.2 ชนิดของความน่าจะเป็น

1) ความน่าจะเป็นเชิงคลาสสิก (classical probability) หรือบางที่เรียกว่าความน่าจะเป็นก่อน (prior probability) คือ ความน่าจะเป็นที่เกิดจากการทดลองสุ่มโดยที่โอกาสในการเกิดเหตุการณ์ต่างๆ จะขึ้นอยู่กับ การทดลอง ดังนั้นความน่าจะเป็นของเหตุการณ์ที่สนใจใดๆ จะเทียบกับผลลัพธ์ทั้งหมดที่เป็นไปได้ เช่น

$$P(A) = \frac{\text{จำนวนผลลัพธ์ในเหตุการณ์ } A}{\text{จำนวนผลลัพธ์ที่เป็นไปได้ทั้งหมด}} = \frac{n(A)}{n(\Omega)}$$

ตัวอย่างที่ 1.4 กล่องใบหนึ่งบรรจุลูกบอล 8 ลูก เป็นสีแดง 3 ลูก และสีดำ 5 ลูก จงหาความน่าจะเป็น

- 1) สุ่มหยิบลูกบอล 1 ลูก และได้สีแดง
- 2) สุ่มหยิบลูกบอล 4 ลูกพร้อมกัน และได้สีแดง 2 ลูก

แนวทางอธิบาย

1) จงหาความน่าจะเป็นสุ่มหยิบลูกบอล 1 ลูก และได้สีแดง
กำหนดให้ $n(\Omega)$ แทนจำนวนผลลัพธ์ที่เป็นไปได้ทั้งหมด นั่นคือ ลูกบอลมี 8 ลูกสุ่มมา 1 ลูก จะได้ $n(\Omega) = 8$ และ A แทนจำนวนผลลัพธ์ที่ได้สีแดง นั่นคือ $n(A) = 3$

$$\text{ดังนั้น } P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{8} = 0.3750$$

2) จงหาความน่าจะเป็นสุ่มหยิบลูกบอล 4 ลูกพร้อม และได้สีแดง 2 ลูก
กำหนดให้ $n(\Omega)$ แทนจำนวนผลลัพธ์ที่เป็นไปได้ทั้งหมดในการสุ่มบอล 4 ลูกพร้อมกันจากบอล 8 ลูก นั่นคือ $n(\Omega) = (8 \times 7 \times 6 \times 5) / (4 \times 3 \times 2 \times 1) = 70$ และ A แทนจำนวนผลลัพธ์ที่ได้สีแดง 2 ลูก นั่นคือ $n(A) = [(3 \times 2 \times 1) / 2] [(5 \times 4) / 2] = 30$

$$\text{ดังนั้น } P(A) = \frac{n(A)}{n(\Omega)} = \frac{30}{70} = 0.4286$$

2) ความน่าจะเป็นเชิงประจักษ์ (empirical probability) หรือ ความถี่สัมพัทธ์ คือ ความน่าจะเป็นที่ใช้ในการประเมินจากการสังเกตจำนวนมาก ดังนั้นความน่าจะเป็นเชิงประจักษ์เป็นการหาความถี่สัมพัทธ์ของแต่ละเหตุการณ์ในการสังเกตนั้นๆ สามารถคำนวณได้ ดังนี้

$$P(A) = \frac{\text{จำนวนผลลัพธ์ในเหตุการณ์ที่สนใจ } A}{\text{จำนวนผลลัพธ์ทั้งหมด}}$$

ตัวอย่างที่ 1.5 ร้านสะดวกซื้อแห่งหนึ่งต้องการสำรวจจำนวนลูกค้าที่เข้ามาซื้อของที่ร้านในแต่ละวัน ในช่วงระยะเวลา 4 สัปดาห์ บันทึกข้อมูลได้ดังนี้

วัน	อาทิตย์	จันทร์	อังคาร	พุธ	พฤหัสบดี	ศุกร์	เสาร์	รวม
จำนวนลูกค้า (คน)	102	45	50	60	68	250	108	683

จงหาความน่าจะเป็นที่จะมีลูกค้ามาซื้อของที่ร้านในวันศุกร์

แนวทางอธิบาย กำหนดให้ A แทนจำนวนลูกค้าที่เข้ามาซื้อของที่ร้านในวันศุกร์

$$P(A) = \frac{\text{จำนวนลูกค้าที่เข้ามาซื้อของที่ร้านในวันศุกร์}}{\text{จำนวนลูกค้าที่เข้ามาซื้อของที่ร้านทั้งหมด}} = \frac{250}{683} = 0.3660$$

3) ความน่าจะเป็นเชิงจิตวิสัย (subjective probability) เป็นการหาความน่าจะเป็นโดยอาศัยการตัดสินใจของตัวบุคคลหรือกลุ่มบุคคลในการหาโอกาสที่จะเกิดขึ้นของเหตุการณ์ที่สนใจศึกษา ซึ่งถูกประเมินโดยความรู้ ประสบการณ์และสัญชาตญาณ ซึ่งไม่มีกฎในการคำนวณความน่าจะเป็นของเหตุการณ์ ตัวอย่างเช่น ร้านค้าปลีกระบุว่าถ้าสถานการณ์ปัจจุบันเป็นเช่นนี้โอกาสที่จะขายสินค้าบางอย่างหมดภายใน 2 วันนี้เป็น 90% หรือนายกค้าทองคำแห่งประเทศไทยคาดการณ์ว่าถ้าสภาพเศรษฐกิจทั่วโลกเป็นเช่นนี้จะทำให้ราคาทองคำปรับลงร้อยละ 15 เป็นต้น การอนุมานเชิงได้

การอนุมานเชิงสถิติจะประยุกต์ใช้ความน่าจะเป็นเชิงคลาสสิกกับการทดลองสุ่มใด ๆ และอาศัยแนวคิดของความน่าจะเป็นเชิงจิตวิสัยมาประยุกต์ใช้ในการอนุมานแบบเบสที่จะต้องอาศัยสารสนเทศที่สำคัญจากสิ่งที่สนใจศึกษามาช่วยในการอนุมานให้มีความแม่นยำมากยิ่งขึ้น ส่วนความน่าจะเป็นเชิงประจักษ์ สำหรับการวิจัยเชิงสำรวจด้วยตัวอย่าง ซึ่งการหาความน่าจะเป็นจะมีหลากหลายวิธีมาช่วยในการหาความน่าจะเป็น ตัวอย่างเช่นการประยุกต์ใช้เทคนิคการนับเพื่อความสะดวกในการหาจำนวนปริภูมิตัวอย่างและจำนวนผลลัพธ์ของเหตุการณ์ที่สนใจ ดังนี้

1. กฎของการคูณ (law of multiplication) ถ้าเหตุการณ์หนึ่ง สามารถแยกเป็น $n_1, n_2, \dots, n_k$ เหตุการณ์ย่อย ดังนั้น จำนวนวิธีทั้งหมดเท่ากับ $n_1 \times n_2 \times n_3 \times \dots \times n_k$ เช่น การโยนเหรียญเที่ยงตรง 1 เหรียญ จำนวน 2 ครั้ง สามารถหาจำนวนปริภูมิตัวอย่าง ได้ดังนี้

การโยนเหรียญครั้งที่ 1 ผลลัพธ์ที่จะเกิดได้ 2 อย่าง คือ {H, T}

การโยนเหรียญครั้งที่ 2 ผลลัพธ์ที่จะเกิดได้ 2 อย่าง คือ {H, T}

ดังนั้น จำนวนปริภูมิตัวอย่างเท่ากับ $2 \times 2 = 4$

2. การเรียงสับเปลี่ยน (permutation) คือ วิธีการเรียงสับเปลี่ยนสิ่งของที่กำหนดให้เป็นลำดับโดยนำมาจัดเรียงทั้งหมด หรือเพียงบางส่วนก็ได้ นั่นคือ ถ้ามีสิ่งของ n ที่แตกต่างกัน นำมาเรียงสับเปลี่ยน r สิ่ง สามารถพิจารณาได้ด้วยเทคนิคการจัดลำดับ ดังนี้

ตำแหน่งที่ 1 เลือกสิ่งของมาวางเรียงได้ n สิ่ง

ตำแหน่งที่ 2 เลือกสิ่งของมาวางเรียงได้ $n-1$ สิ่ง

ตำแหน่งที่ 3 เลือกสิ่งของมาวางเรียงได้ $n-2$ สิ่ง

ตำแหน่งที่ r เลือกสิ่งของมาวางเรียงได้ $n-(r-1) = n-r+1$ สิ่ง

ดังนั้น จำนวนวิธีการเรียงสับเปลี่ยนทั้งหมด $= n(n-1)(n-2)\dots(n-r+1)$

$$= n(n-1)(n-2)\dots(n-r+1) \frac{(n-r)(n-r-1)\dots 3 \times 2 \times 1}{(n-r)(n-r-1)\dots 3 \times 2 \times 1}$$

$$= \frac{n!}{(n-r)!} = {}^n P_r$$

$$\text{นั่นคือ จำนวนวิธีการเรียงสับเปลี่ยน} = {}^n P_r = \frac{n!}{(n-r)!}$$

n แฟกทอเรียล (factorial) แทนด้วย $n!$
 $n! = n(n-1)(n-2) \dots 3 \times 2 \times 1$
 $0! = 1, 1! = 1$

ตัวอย่างที่ 1.5 ในการเลือกตัวแทนของนักศึกษาห้องหนึ่งจากนักศึกษาทั้งหมด 5 คน จงหาจำนวนวิธีที่จะเลือกได้ทั้งหมด ในการเลือกตำแหน่งหัวหน้าห้องและรองหัวหน้าห้อง

แนวทางอธิบาย จากผู้สมัครทั้งหมด 5 คน จะได้ $n = 5$

นำมาจัดลำดับ คือ ตำแหน่งหัวหน้าห้องและรองหัวหน้าห้อง จะได้ $r = 2$

$$\text{จำนวนวิธีการเรียงสับเปลี่ยน } {}^n P_r = \frac{n!}{(n-r)!}$$

$${}^5 P_2 = \frac{5!}{(5-2)!} = \frac{5!}{3!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1} = 5 \times 4 = 20 \quad \text{วิธี}$$

นั่นคือ จำนวนวิธีทั้งหมดที่จะเลือกตัวแทนของนักศึกษาแต่ละคนอยู่ในตำแหน่งหัวหน้าห้องและรองหัวหน้าห้องเท่ากับ 20 วิธี จากนักศึกษาทั้งหมด 5 คน