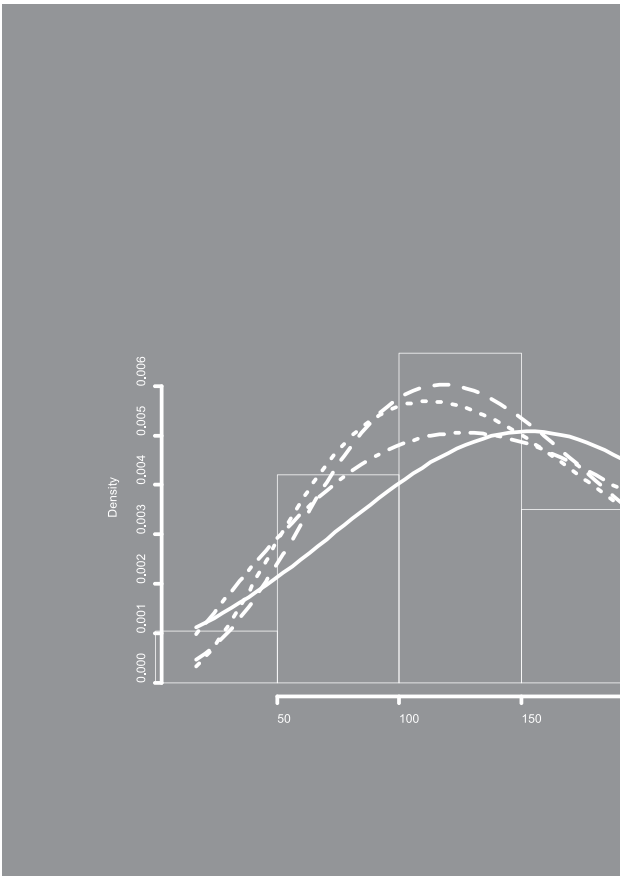
ความน่าจะเป็น
กับการประยุกต์ใช้ในการอนุมานเชิงสถิติ

Probability with Application to Statistical Inference

ฉบับปรับปรุง

มานัดถุ์ คำกอง
ภาควิชาสถิติ คณะวิทยาศาสตร์
มหาวิทยาลัยเชียงใหม่



ความน่าจะเป็น
กับการประยุกต์ใช้ในการอนุมานเชิงสถิติ

Probability with Application to Statistical Inference

มานัดถุ์ คำกอง

ภาควิชาสถิติ คณะวิทยาศาสตร์

มหาวิทยาลัยเชียงใหม่

ฉบับปรับปรุง 2563

คำนำ (ฉบับปรับปรุง)

หนังสือความน่าจะเป็นกับการประยุกต์ใช้ในการอนุมานเชิงสถิติ (ฉบับปรับปรุง) ผู้เขียนได้ปรับปรุงจากเล่มเดิมโดยการเพิ่มเนื้อหาในส่วนขององค์ประกอบของการอนุมานเชิงสถิติ แยกเป็นหลักการเบื้องต้นของการลดมิติข้อมูล การประมาณค่าพารามิเตอร์ และการทดสอบสมมุติฐาน เพื่อให้หนังสือเล่มนี้ได้ตอบสนองต่อการนำแนวความคิดของการเกิดทฤษฎีการแจกแจงทางสถิติ ถูกนำไปการประยุกต์ใช้ในการแก้ไขปัญหาการอนุมานเชิงสถิติขั้นสูงต่อไป

ผู้เขียนหวังว่าหนังสือเล่มนี้จะเป็นประโยชน์แก่นักศึกษาและผู้ที่สนใจศึกษาที่มีความรู้เบื้องต้นทางคณิตศาสตร์ที่เกี่ยวข้องกับทางแคลคูลัสและพีชคณิตเบื้องต้นมาบ้าง จะได้เห็นประโยชน์ของการประยุกต์ใช้ความน่าจะเป็นต่อการอนุมานเชิงสถิติ

ผู้เขียนกราบขอบพระคุณ คุณพ่อสีผู้ที่ล่วงลับไปแล้วในระหว่างที่ผู้เขียนอยู่ในช่วงปรับปรุงหนังสือเล่มนี้ คุณพ่อสีเป็นแบบอย่างที่ดี และคอยเป็นสิ่งยึดเหนียวจิตใจให้ผู้เขียนมีความอดทนและมานะ สู้ต่อปัญหา/อุปสรรคต่าง ๆ คุณแม่ผู้เป็นแรงใจที่ดีในการปรับปรุงหนังสือ รวมทั้งภรรยาและลูกชายทั้งสองที่สนับสนุนให้กำลังใจและส่งเสริมการผลิตหนังสือของผู้เขียนมาโดยตลอด กราบขอบพระคุณ ครู-อาจารย์ทุกท่านที่ได้ให้ความรู้แก่ผู้เขียนตลอดมา กราบขอบพระคุณหลักธรรมคำสั่งสอนขององค์สมเด็จพระสัมมาสัมพุทธเจ้าศาสดาของพระพุทธศาสนาที่ผู้เขียนใช้เป็นเครื่องยึดเหนียวจิตใจและนำมาปฏิบัติในชีวิตตลอดเรื่อยมา และขอขอบคุณภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ ที่ให้การสนับสนุนและเอื้อประโยชน์ให้ผู้เขียนได้ผลิตหนังสือเล่มนี้

ผู้เขียนต้องกราบขอบคุณผู้ประเมินหนังสือเล่มนี้ที่ได้ให้ข้อเสนอแนะที่เป็นประโยชน์ต่อการแก้ไขหนังสือเล่มนี้ให้สมบูรณ์มากยิ่งขึ้น และขอบคุณ คุณธนโชติ ไชยโต นักศึกษาระดับปริญญาเอก สาขาสถิติประยุกต์ ที่ช่วยในการจัดพิมพ์แก้ไขหนังสือเล่มนี้ อีกทั้งยังขอบพระคุณสำนักพิมพ์มหาวิทยาลัยเชียงใหม่ ที่ได้ให้ความกรุณาอนุญาตให้ผู้เขียนปรับปรุงหนังสือเล่มนี้

ท้ายที่สุดผู้เขียนก็เป็นปุถุชนคนหนึ่งที่มีความไม่สมบูรณ์ หากหนังสือเล่มนี้มีข้อบกพร่องหรือความผิดพลาดประการใด โปรดแจ้งกลับมาที่ผู้เขียนที่ e-mail: manad.k@cmu.ac.th ผู้เขียนกราบขออภัยเป็นอย่างสูงในข้อบกพร่อง/ความผิดพลาดนั้น และจะได้นำมาปรับแก้ไขในโอกาสต่อไป

มานัดถุ์ คำกอง
สิงหาคม 2563

ตัวอย่าง

ตัวอย่าง

คำนำ

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

ผู้เขียนหวังว่าหนังสือเล่มนี้จะเป็นประโยชน์แก่นักศึกษาและผู้ที่มีความสนใจศึกษาในบางเรื่องที่สนใจศึกษาการประยุกต์ใช้ความน่าจะเป็นต่อการอนุมานเชิงสถิติ ทั้งนี้ผู้เขียนได้เล็งเห็นถึงความสำคัญในการใช้โปรแกรมสำเร็จรูปเพื่อช่วยในการวิเคราะห์ปัญหาด้านสถิติ ที่สามารถเข้าถึงได้ง่ายและสามารถพัฒนาการเขียนโปรแกรมเพื่อตอบสนองการวิเคราะห์ข้อมูลได้

ผู้เขียนขอกราบขอบพระคุณพ่อ-แม่ พี่ ๆ น้อง รวมทั้งภรรยาและลูกชายที่สนับสนุนให้กำลังใจและส่งเสริมการผลิตหนังสือของผู้เขียนมาโดยตลอด กราบขอบพระคุณ ครู-อาจารย์ทุกท่านที่ได้ให้ความรู้แก่ผู้เขียนตลอดมา กราบขอบพระคุณหลักธรรมคำสั่งสอนขององค์สมเด็จ พระสัมมาสัมพุทธเจ้าศาสดาของพระพุทธศาสนาที่ผู้เขียนใช้เป็นเครื่องยึดเหนียวจิตใจและนำมาปฏิบัติในชีวิตตลอดเรื่อยมา และขอขอบคุณภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ ที่ให้การสนับสนุนและเอื้อประโยชน์ให้ผู้เขียนได้ผลิตหนังสือเล่มนี้

ผู้เขียนต้องขอบคุณผู้ประเมินหนังสือเล่มนี้ที่ได้ให้ข้อเสนอแนะที่เป็นประโยชน์ต่อการแก้ไขหนังสือเล่มนี้ให้สมบูรณ์มากยิ่งขึ้น อีกทั้งยังขอบพระคุณสำนักพิมพ์มหาวิทยาลัยเชียงใหม่ ที่ได้ให้ความกรุณาในการจัดพิมพ์หนังสือเล่มนี้

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

มานัดถุ์ คำกอง
มกราคม 2560

ร่างต้นฉบับ

ตัวอย่าง

สารบัญ

คำนำ (ฉบับปรับปรุง) III

คำนำ V

สารบัญ VII

สารบัญตาราง XII

สารบัญภาพ XIII

บทที่ 1 ความน่าจะเป็นและสถิติ 1

1.1 บทนำ 2

1.2 ความน่าจะเป็นและเทคนิคการนับ 9

1.3 แนวคิดพื้นฐานความน่าจะเป็นและสมบัติความน่าจะเป็น 18

แบบฝึกหัด 29

บรรณานุกรม 32

บทที่ 2 ตัวแปรสุ่ม ฟังก์ชันความน่าจะเป็นและสมบัติที่สำคัญ 33

2.1 บทนำ 34

2.2 ตัวแปรสุ่มและฟังก์ชันความน่าจะเป็น 35

2.3 การคาดหมายและความแปรปรวน 44

2.4 โมเมนต์และฟังก์ชันก่อกำเนิดโมเมนต์ 55

แบบฝึกหัด 65

บรรณานุกรม 72

ตัวอย่าง

สารบัญ

บทที่ 3 ฟังก์ชันความน่าจะเป็นร่วมและฟังก์ชันความน่าจะเป็นมีเงื่อนไข 73

3.1 บทนำ 74
3.2 ฟังก์ชันความน่าจะเป็นร่วม 75
3.3 ฟังก์ชันความน่าจะเป็นมีเงื่อนไข 90
3.4 การคาดหมายมีเงื่อนไขและความแปรปรวนมีเงื่อนไข 98
แบบฝึกหัด 109
บรรณานุกรม 114

บทที่ 4 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มไม่ต่อเนื่อง 115

4.1 บทนำ 116
4.2 การแจกแจงเอกรูปไม่ต่อเนื่อง 117
4.3 การแจกแจงแบร์นูลลีและการแจกแจงทวินาม 119
4.4 การแจกแจงเรขาคณิตและการแจกแจงทวินามเชิงลบ 128
4.5 การแจกแจงไฮเพอร์จีออเมตริก 140
4.6 การแจกแจงปัวซง 144
4.7 ความสัมพันธ์ระหว่างการแจกแจงไม่ต่อเนื่อง 149
แบบฝึกหัด 160
บรรณานุกรม 166

ตัวอย่าง

สารบัญ

บทที่ 5 การแจกแจงความน่าจะเป็นของตัวแปรสุ่มต่อเนื่อง 167

5.1 บทนำ 168
5.2 การแจกแจงเอกรูป 169
5.3 การแจกแจงแบบเลขชี้กำลังและการแจกแจงแกมมา 173
5.4 การแจกแจงบีตา 188
5.5 การแจกแจงปรกติ 192
5.6 ความสัมพันธ์ระหว่างการแจกแจงต่อเนื่อง 203
แบบฝึกหัด 205
บรรณานุกรม 210

บทที่ 6 การแจกแจงความน่าจะเป็นของฟังก์ชันตัวแปรสุ่ม 211

6.1 บทนำ 212
6.2 เทคนิคฟังก์ชันความน่าจะเป็นสะสม 213
6.3 เทคนิคการเปลี่ยนตัวแปร 222
6.4 เทคนิคฟังก์ชันก่อกำเนิดโมเมนต์ 237
6.5 การแจกแจงของฟังก์ชันตัวแปรสุ่มปรกติ 241
แบบฝึกหัด 257
บรรณานุกรม 263

ตัวอย่าง

สารบัญ

บทที่ 7 ทฤษฎีบทขีดจำกัดเบื้องต้น 265

7.1 บทนำ 266
7.2 แนวคิดการลู่เข้าเบื้องต้น 267
7.3 กฎจำนวนมาก 275
7.4 ทฤษฎีบทขีดจำกัดส่วนกลาง 281
แบบฝึกหัด 294
บรรณานุกรม 298

บทที่ 8 หลักการเบื้องต้นของการลดมิติข้อมูล 299

8.1 บทนำ 300
8.2 แนวคิดเบื้องต้นเกี่ยวกับตัวอย่างสุ่ม 301
8.3 การแจกแจงของสถิติอันดับ 311
8.4 หลักการความพอเพียงและการใช้ประโยชน์ 323
8.5 หลักการภาวะน่าจะเป็นและการใช้ประโยชน์ 331
แบบฝึกหัด 342
บรรณานุกรม 346

บทที่ 9 การประมาณค่าพารามิเตอร์ 347

9.1 บทนำ 348
9.2 การประมาณค่าแบบจุดและสมบัติที่สำคัญ 349
9.3 การประมาณค่าแบบช่วงและสมบัติที่สำคัญ 382
9.4 การกำหนดขนาดตัวอย่าง 399
แบบฝึกหัด 403
บรรณานุกรม 408

ตัวอย่าง

สารบัญ

บทที่ 10 การทดสอบสมมุติฐาน 409

10.1 บทนำ 410
10.2 สมมุติฐานทางสถิติและการประเมินผลการทดสอบ 411
10.3 วิธีการหาสถิติทดสอบ 420
10.4 วิธีของการประเมินผลการทดสอบ 447
แบบฝึกหัด 450
บรรณานุกรม 454

ภาคผนวก 455

ภาคผนวก 1 ทบทวนความรู้เชิงคณิตศาสตร์ 456
ภาคผนวก 2 โปรแกรมอาร์เบื้องต้น 460
ภาคผนวก 3 การสร้างตารางความน่าจะเป็นของการแจกแจงทวินาม 465
ภาคผนวก 4 การสร้างตารางความน่าจะเป็นของการแจกแจงปัวซง 466
ภาคผนวก 5 การสร้างตารางการแจกแจงสะสมปรกติมาตรฐาน 467
ภาคผนวก 6 การสร้างตารางการแจกแจงสะสมที 468

แนวทางอธิบายแบบฝึกหัด 471

ดัชนี 505

ตัวอย่าง

สารบัญตาราง

ตารางที่ 4.1 ค่าคุณลักษณะที่สำคัญของการแจกแจงปัวซงและการแจกแจงไม่ต่อเนื่องที่เกี่ยวข้อง 53

ตารางที่ 10.1 ความสัมพันธ์ระหว่างความผิดพลาดแบบที่ 1 และ 2 414

ตัวอย่าง

สารบัญภาพ

ภาพที่ 1.1	ขอบข่ายของการศึกษาสถิติศาสตร์	4
ภาพที่ 1.2	ชนิดของความน่าจะเป็น	10
ภาพที่ 2.1	ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม X ที่สนใจศึกษา	40
ภาพที่ 2.2	ฟังก์ชันการแจกแจงสะสมของตัวแปรสุ่ม Y ที่สนใจศึกษา	41
ภาพที่ 2.3	ลักษณะโค้งความถี่ของตัวแปรสุ่มที่วัดด้วยค่าสัมประสิทธิ์ความเบ้	56
ภาพที่ 2.4	ลักษณะโค้งความถี่ของตัวแปรสุ่มที่วัดด้วยค่าสัมประสิทธิ์ความโด่ง	57
ภาพที่ 3.1	แนวความคิดของการใช้ทฤษฎีบทของเบส์	94
ภาพที่ 4.1	ฟังก์ชันความน่าจะเป็นของการแจกแจงเอกรูปไม่ต่อเนื่อง เมื่อ N = 4, 6 และ 10	118
ภาพที่ 4.2	ฟังก์ชันความน่าจะเป็นของการแจกแจงทวินาม เมื่อ n = 5, 15, 30 และ ค่า p = 0.1, 0.5, 0.9	123
ภาพที่ 4.3	ฟังก์ชันความน่าจะเป็นของการแจกแจงเรขาคณิต เมื่อ p = 0.2, 0.5 และ 0.8	131
ภาพที่ 4.4	ฟังก์ชันความน่าจะเป็นการแจกแจงทวินามเชิงลบ เมื่อ r = 10, 20 และ p = 0.15, 0.5, 0.85	139
ภาพที่ 4.5	ฟังก์ชันความน่าจะเป็นของการแจกแจงแจงไฮเพอร์จีออเมตริก M =10, $N-M$ =10, 20 และ n = 3, 5, 7	143
ภาพที่ 4.6	ฟังก์ชันความน่าจะเป็นของการแจกแจงปัวซง เมื่อ λ = 3, 10 และ 15	146
ภาพที่ 4.7	ความสัมพันธ์ระหว่างการแจกแจงปัวซงและการแจกแจงไม่ต่อเนื่องที่เกี่ยวข้อง	155
ภาพที่ 5.1	ฟังก์ชันความน่าจะเป็นของการแจกแจงเอกรูปชนิดต่อเนื่อง	171
ภาพที่ 5.2	ฟังก์ชันความน่าจะเป็นของการแจกแจงแบบเลขชี้กำลังด้วยพารามิเตอร์ λ = 1, 2 และ 3	176

ตัวอย่าง

สารบัญภาพ

ภาพที่ 5.3 ฟังก์ชันความน่าจะเป็นของการแจกแจงแกมมาด้วยพารามิเตอร์ α =1, 2, 3 และ λ =1 182

ภาพที่ 5.4 ฟังก์ชันความน่าจะเป็นของการแจกแจงไคกำลังสองด้วยองศาเสรี ν =1, 2 และ 30 187

ภาพที่ 5.5 ฟังก์ชันความน่าจะเป็นของการแจกแจงบีตาด้วยพารามิเตอร์ a และ b 191

ภาพที่ 5.6 ฟังก์ชันความน่าจะเป็นของการแจกแจงปรกติด้วยค่าเฉลี่ย 1 และส่วนเบี่ยงเบนมาตรฐาน 1, 3 และ 6 196

ภาพที่ 5.7 พื้นที่ใต้ส่วนโค้งการแจกแจงปรกติที่ห่างจากค่าเฉลี่ยเป็นระยะ σ, σ 2 และ σ 3 197

ภาพที่ 5.8 การแจกแจงปรกติสองตัวแปรด้วยค่าเฉลี่ย $\mu_1 = \mu_2$ =0 ความแปรปรวน $\sigma_{11} = \sigma_{22} = 10$ และค่าสัมประสิทธิ์สหสัมพันธ์ σ = 0 และ σ = 0.75 202

ภาพที่ 5.9 ความสัมพันธ์ระหว่างการแจกแจงแกมมาและการแจกแจงต่อเนื่องที่เกี่ยวข้อง 204

ภาพที่ 6.1 การสร้างการแจงแจงความน่าจะเป็นของฟังก์ชันตัวแปรสุ่ม $Y = g(X)$ 212

ภาพที่ 6.2 ฟังก์ชันการแจกแจงสะสมของ $Y = g(X)$ กรณี $g(.)$ เป็นฟังก์ชันต่อเนื่องลดโดยแท้จริง 224

ภาพที่ 6.3 ฟังก์ชันการแจกแจงสะสมของ $Y = g(X)$ กรณี $g(.)$ เป็นฟังก์ชันต่อเนื่องลดโดยแท้จริง 244

ภาพที่ 6.4 ฟังก์ชันความน่าจะเป็นของ $f(x) = \frac{x^2}{2}, \ -1 < x < 3$ 228

ภาพที่ 6.5 ฟังก์ชันความน่าจะเป็นของการแจกแจงทีและการแจกแจงปรกติมาตรฐาน 252

ภาพที่ 6.6 ฟังก์ชันความน่าจะเป็นของการแจกแจงเอฟ 256

ภาพที่ 10.1 การเลือกสถิติทดสอบค่าเฉลี่ย 1 ประชากรของการแจกแจงปรกติ 440

ตัวอย่าง

ตัวอย่าง

01

ความน่าจะเป็นและสถิติ

PROBABILITY AND STATISTICS

PROBABILITY WITH APPLICATION TO STATISTICAL INFERENCE

1.1 บทนำ (Introduction)

ความน่าจะเป็นและสถิติคืออะไร? ทำไมต้องศึกษาความน่าจะเป็น?

ในชีวิตประจำวันของแต่ละบุคคลมีโอกาสได้รับข้อมูลและข่าวสารต่าง ๆ อย่างมากมาย ตัวอย่างเช่น การรับฟังรายงานพยากรณ์อากาศประจำวันจากกรมอุตุนิยมวิทยาเพื่อให้ได้รับรู้ข้อมูลสภาพภูมิอากาศเพื่อช่วยในการวางแผนด้านการเดินทาง การรับฟังราคาซื้อขายทองคำต่อวันจากสมาคมค้าทองแห่งประเทศไทยเพื่อให้เห็นแนวโน้มของราคาทองคำ เป็นต้น ดังนั้น จำเป็นต้องมีการวิเคราะห์ข้อมูลที่ได้รับข่าวสารมาเพื่อช่วยในการตัดสินใจในหลาย ๆ ด้านไม่ว่าจะช่วยในการดำรงชีวิตประจำวัน การตัดสินใจที่ประกอบการลงทุนต่าง ๆ ซึ่งต้องอาศัยการศึกษาศาสตร์ที่เกี่ยวข้องกับสถิติ โดยที่สถิติศาสตร์ (statistics) เป็นวิทยาศาสตร์ที่เกี่ยวข้องกับการวางแผนการศึกษา การเก็บรวบรวมข้อมูล (หรือการสร้างแบบจำลอง) และการวิเคราะห์ข้อมูลเพื่อใช้ประโยชน์ในการตัดสินใจและการค้นพบองค์ความรู้ทางวิทยาศาสตร์ภายใต้ข้อจำกัดด้านข้อมูลและตัวแปรที่ศึกษา นั่นคือ **สถิติศาสตร์ เป็นวิทยาศาสตร์ของการเรียนรู้จากข้อมูล** (Ott และ Longnecker, 2010: 2) การอนุมานเชิงสถิติอาศัยสารสนเทศที่ได้จากตัวอย่างสุ่ม $X_1, X_2, ..., X_n$ ที่สุ่มมาศึกษาขนาด n เพื่อสรุปเกี่ยวกับค่าพารามิเตอร์ที่ไม่ทราบค่า θ คือ คาดการณ์ว่าตัวอย่างสุ่มได้เก็บข้อมูลสารสนเทศที่สำคัญเกี่ยวกับพารามิเตอร์ที่สนใจศึกษาไว้หมดแล้ว ซึ่งภายใต้การสุ่มตัวอย่างมาศึกษาที่ถูกสุ่มมาจากประชากรที่สนใจศึกษาที่มีฟังก์ชันความน่าจะเป็น $f_X(x|\theta)$ โดยทั่วไปจะไม่ทราบรูปแบบฟังก์ชันความน่าจะเป็นที่แท้จริง แต่จะมีการสมมุติว่าประชากร ที่สนใจศึกษามีลักษณะรูปแบบฟังก์ชันความน่าจะเป็นลักษณะอย่างใดอย่างหนึ่งที่มีพารามิเตอร์ที่สนใจศึกษาไม่ทราบค่า θ ซึ่งจะได้ศึกษาอย่างละเอียดในบทถัดไป หากในกระบวนการอนุมานเชิงสถิติที่จำเป็นต้องระบุการแจกแจงของตัวอย่างสุ่มที่มีค่าพารามิเตอร์ที่สนใจศึกษากำกับอยู่เรียกการอนุมานนี้ว่า สถิติศาสตร์อิงพารามิเตอร์

ตัวอย่าง

(parametric statistics) และหากการอนุมานนั้นไม่จำเป็นต้องระบุการแจกแจงของตัวอย่างสุ่ม จะเรียกการอนุมานนั้นว่า สถิติศาสตร์ไม่อิงพารามิเตอร์ (non-parametric statistics) กรณีศึกษาของผู้จัดการฝ่ายขายโทรศัพท์มือถือยี่ห้อหนึ่ง ต้องการศึกษาสัดส่วนของเครื่องโทรศัพท์มือถือยี่ห้อนี้ที่จำหน่ายให้ผู้บริโภคจะเสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปีจะมีมากน้อยเพียงใด ดังนั้นในการศึกษานี้จะสนใจประมาณค่าพารามิเตอร์ คือ สัดส่วน (proportion, p) ของเครื่องโทรศัพท์มือถือของบริษัทจะเสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี ในการศึกษานี้ผู้จัดการฝ่ายขายอาจสุ่มข้อมูลที่ขายเครื่องโทรศัพท์มือถือยี่ห้อนี้จำนวน n มาศึกษา นั่นคือ

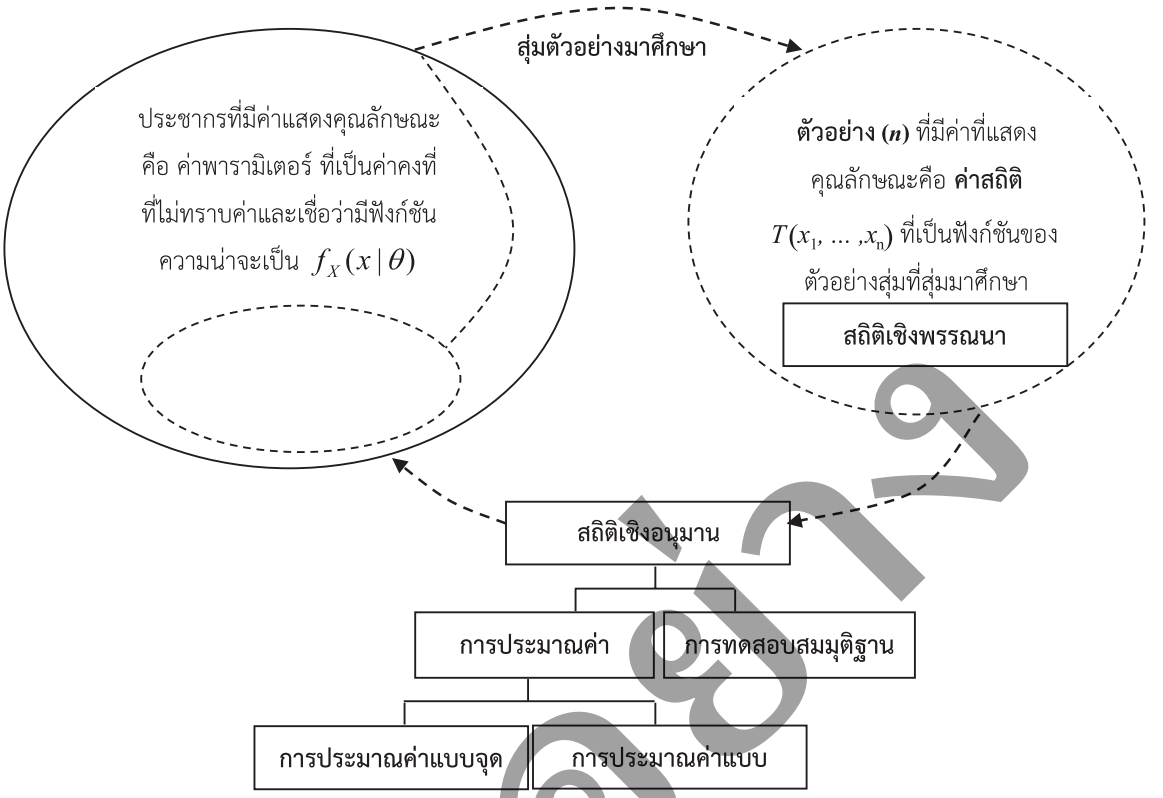
$$X=\begin{cases} 0 \text{ แทนโทรศัพท์มือถือที่}\textbf{ไม่เสีย}\text{ก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี ด้วยความน่าจะเป็น } 1-p \\ 1 \text{ แทนโทรศัพท์มือถือที่}\textbf{เสีย}\text{ก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี ด้วยความน่าจะเป็น } p \end{cases}$$

จากข้อมูลที่สุ่มมาศึกษาทางผู้จัดการฝ่ายขายจะไม่ทราบลักษณะฟังก์ชันความน่าจะเป็น $f_X(x \mid p)$ ที่แท้จริงของจำนวนเครื่องโทรศัพท์มือถือที่เสียก่อนครบระยะเวลาที่ทางบริษัทรับประกัน 1 ปี จากลักษณะของการเกิดของ X แล้ว สามารถเขียนรูปแบบฟังก์ชันความน่าจะเป็นของการเกิดขึ้นของ X ภายใต้พารามิเตอร์ที่สนใจศึกษา p ได้ ดังนี้

$$f(X=x \mid p)=p^x(1-p)^{1-x}, \quad x=0,1; \quad 0\le p\le 1$$

และศูนย์ที่อื่น ๆ โดยที่ไม่ทราบค่าพารามิเตอร์ p ที่สนใจศึกษา แต่ทราบขอบเขตที่เป็นไปได้ทั้งหมดของ p คือ $p=\{p:0\le p\le 1\}$ ซึ่งการอนุมานนี้ได้มีการกำหนดฟังก์ชันความน่าจะเป็นของการเกิดขึ้นของสิ่งที่สนใจศึกษาที่มีพารามิเตอร์ไม่ทราบค่า p ในการอนุมานค่าพารามิเตอร์ ที่ไม่ทราบค่าสามารถศึกษาผ่านขอบข่ายของการศึกษาสถิติศาสตร์ ดังภาพที่ 1.1

ตัวอย่าง



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 จำนวนคำที่ปรากฏในหนังสือความน่าจะเป็นกับการประยุกต์ใช้ในการอนุมานเชิงสถิติ เป็นต้น

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

ตัวอย่าง

2) ตัวอย่าง (sample) หมายถึง ชุดขององค์ประกอบหรือขอบเขตบางส่วนของประชากรที่สนใจศึกษา โดยการเลือกมาศึกษาเพื่อเป็นตัวแทนในการหาข้อสรุปเกี่ยวกับประชากรนั้น ๆ

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

4) ค่าสถิติ (statistic) หมายถึง ฟังก์ชันของค่าสังเกตที่วัดจากหน่วยต่าง ๆ ที่ถูกเลือกขึ้นมาเป็นตัวอย่าง ***ซึ่งจะไม่มีค่าพารามิเตอร์อื่น ๆ ที่ไม่ทราบค่าปนอยู่*** ค่าสถิติที่ได้จะใช้เป็นตัวประมาณค่าของพารามิเตอร์ เช่น จากการทดลองปลูกข้าวด้วยวิธีการใหม่ของเกษตรกรอำเภอท่านหนึ่งในพื้นที่ต่าง ๆ อย่างละ 1 ไร่ จำนวนทั้งหมด 30 ไร่ เป็นตัวอย่างในการศึกษาภายใต้การควบคุมปัจจัยต่าง ๆ ให้เหมือนกัน เมื่อเสร็จสิ้นการทดลองได้ทำการวัดผลผลิตแต่ละไร่โดยการหาค่าเฉลี่ยของตัวอย่าง (sample mean) $\bar{X}$ และความแปรปรวนของตัวอย่าง (sample variance) S^2

5) หน่วยตัวอย่าง (sampling unit) หมายถึง หน่วยที่ต้องการศึกษาหรือจัดรายละเอียดข้อเท็จจริงบางประการ อาจจะเป็นหน่วยเล็ก ๆ เพียงหน่วยเดียว หรือเป็นกลุ่มของหน่วยเล็ก ๆ หลาย ๆ หน่วยรวมกันก็ได้

6) หน่วยการแจงนับ (enumeration unit) หมายถึง หน่วยหรือสิ่งของที่จะเก็บรวบรวมข้อมูลในบางกรณีมีความหมายคล้ายกับหน่วยตัวอย่าง เช่น การศึกษาค่าใช้จ่ายต่อเดือนของนักศึกษามหาวิทยาลัยเชียงใหม่ ประจำปีการศึกษา 2559 หน่วยตัวอย่าง คือ นักศึกษามหาวิทยาลัยเชียงใหม่ที่มีสภาพการเป็นนักศึกษาแต่ละคนในปีการศึกษา 2559 เนื่องจากการสำรวจสามารถสอบถามค่าใช้จ่ายต่อเดือนของนักศึกษาแต่ละคนได้ ดังนั้น นักศึกษาแต่ละคนเป็นหน่วยการแจงนับด้วย แต่ในบางกรณีที่หน่วยตัวอย่างและหน่วยการแจงนับมีความหมายที่แตกต่างกัน เช่น การสำรวจผลผลิตข้าวของครัวเรือนเกษตรกรในจังหวัดเชียงใหม่ ปี พ.ศ. 2559 โดยกำหนดให้ครัวเรือนเกษตรกรเป็นหน่วยตัวอย่าง ซึ่งครัวเรือนไม่สามารถให้ข้อมูลได้ ดังนั้น จำเป็นต้องถามบุคคลในครัวเรือน เช่น หัวหน้าครัวเรือนแต่ละครัวเรือน กรณีนี้หัวหน้าครัวเรือนเป็นหน่วยการแจงนับ

7) กรอบตัวอย่าง (sampling frame) หมายถึง บัญชีรายชื่อของหน่วยตัวอย่างที่ประกอบด้วยรายละเอียดเกี่ยวกับตำแหน่งที่อยู่ของหน่วยตัวอย่างเหล่านั้น

8) แผนแบบการเลือกตัวอย่าง (sample design) หมายถึง วิธีการเลือกหน่วยตัวอย่างต่าง ๆ เพื่อให้ได้ตัวอย่างที่เป็นตัวแทนของประชากรที่สนใจศึกษา

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

ตัวอย่าง

10) สถิติเชิงอนุมาน (inferential statistics) หมายถึง การใช้ระเบียบวิธีการทางสถิติเพื่อจัดทำข้อสรุปของตัวอย่างที่สุ่มมาศึกษาเพื่ออนุมานไปสู่ประชากรโดยการประเมินความน่าเชื่อถือของวิธีการของสถิติดังกล่าวนั้น ๆ

ตัวอย่างที่ 1.1 นักวิจัยสิ่งแวดล้อมท่านหนึ่งได้รับข้อร้องเรียนจากชาวบ้านเกี่ยวกับโรงงานผลิตสารกำจัดศัตรูพืชแห่งหนึ่งที่ตั้งอยู่ใกล้กับแม่น้ำได้ลักลอบปล่อยน้ำเสียลงสู่แม่น้ำ ทำให้สิ่งมีชีวิต ที่อาศัยอยู่ในแม่น้ำตายเป็นจำนวนมากและแม่น้ำเน่าเสียไม่สามารถอุปโภคได้ นักวิจัยท่านนี้ จึงได้สุ่มตรวจระบบบำบัดน้ำเสียของโรงงานแห่งนี้ปริมาณ 500 มิลลิลิตร เพื่อตรวจวัดว่าค่า pH, BOD, SS, FOG และ TDS จากน้ำเสียที่ปล่อยจากโรงงานแห่งนี้มาจำนวน 30 วัน จะมีค่าเกินกว่าที่คณะกรรมการควบคุมมลพิษกำหนดหรือไม่ 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 มก/ล
• 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 ปริมาณน้ำเสียที่ปล่อยจากโรงงานแหล่งนี้ทั้งหมด

ตัวอย่าง คือ ปริมาณน้ำเสียที่ปล่อยจากโรงงานแหล่งนี้ที่สุ่มมาตรวจในแต่ละวัน ปริมาณ 500 มิลลิลิตร จำนวน 30 วัน

หน่วยตัวอย่าง คือ ปริมาณน้ำเสียที่ปล่อยจากโรงงานแหล่งนี้ในแต่ละวัน

หน่วยการแจงนับ คือ เนื่องจากปริมาณน้ำเสียที่ปล่อยจากโรงงานแหล่งนี้สามารถวัดค่าต่าง ๆ ที่สนใจศึกษาได้ เช่น pH, BOD, SS, FOG และ TDS ดังนั้น หน่วยการแจงนับจะเหมือนกับหน่วยตัวอย่าง

ตัวอย่าง

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

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

จากตัวอย่างที่ 1.1 จะเห็นได้ว่าการใช้ตัวอย่างสุ่มเพียงบางส่วนมาศึกษาเพื่อกลับไปอ้างอิงถึงประชากรทั้งหมดนั้นซึ่งอาจมีความผิดพลาดที่อาจเกิดขึ้นได้ ดังนั้นจะต้องใช้ความระมัดระวังในการสรุปผล ในการทดลองหลาย ๆ สาขาวิชาทางวิทยาศาสตร์ ก็จะทำการศึกษาในลักษณะนี้เช่นกัน แต่จะมีการควบคุมปัจจัยหรือสภาพแวดล้อมในการทดลองตามวัตถุประสงค์ของการศึกษา ซึ่งการอนุมานในลักษณะนี้เรียกว่า**การอนุมานอุปนัย (inductive inference)**

การอนุมานอุปนัยเป็นกระบวนการที่มีความเสี่ยงต่อการสรุปผลการศึกษาอย่างยิ่ง เพราะในความเป็นจริงแล้ว การอนุมานอุปนัยอาศัยตรรกะที่มีความไม่แน่นอนช่วยในการอนุมาน ซึ่งความไม่แน่นอนนั้นสามารถที่จะวัดระดับของความไม่แน่นอนได้ หากการทดลองได้รับ การดำเนินการตามหลักการที่ถูกต้อง สำหรับบทบาทหนึ่งของสถิติที่มีผลต่อเทคนิคการทำข้อสรุปแบบอุปนัย นั่นคือการวัดระดับของความไม่แน่นอนของการวินิจฉัยดังกล่าวด้วยความน่าจะเป็น (probability) ดังนั้นความไม่แน่นอนจะถูกวัดออกมาในแง่ของความน่าจะเป็น นั่นเป็นเหตุผล ที่สำคัญที่ควรมีการศึกษาทฤษฎีของความน่าจะเป็นเพื่อประยุกต์ใช้ในทฤษฎีการอนุมานเชิงสถิติต่อไป

สำหรับการอนุมานอีกแบบหนึ่งที่นิยมใช้ในการพิสูจน์ทฤษฎีบทความเป็นจริงทางคณิตศาสตร์ นั่นคือ **การอนุมานนิรนัย (deductive inference)** โดยที่การอนุมานนิรนัยจะอาศัยเหตุปัจจัยของความเป็นจริงเพื่อเป็นข้อสรุปผลจากเหตุที่ได้

ตัวอย่างที่ 1.2 พิจารณาข้อความจริง 2 ประการ ดังนี้

1) หลักฐานหลัก: มุมภายในหนึ่งมุมของรูปสามเหลี่ยมมุมฉากแต่ละรูปเท่ากับ 90 องศา

2) หลักฐานรอง: สามเหลี่ยม B เป็นรูปสามเหลี่ยมมุมฉาก ถ้ายอมรับหลักฐานทั้งสองอย่างนี้ แล้วจะถูกบังคับให้สรุปว่า

3) ข้อสรุป: หนึ่งในมุมของรูปสามเหลี่ยม B เท่ากับ 90 องศา

ตัวอย่าง

จากตัวอย่างที่ 1.2 จะเห็นได้ว่าเป็นการอนุมานนิรนัยที่ข้อสรุปความจริง 3) จะเกิดจากข้อเท็จจริงที่ยอมรับหลักฐานจาก 1) และ 2) นั่นคือ การอนุมานนิรนัยเป็นสิ่งสำคัญอย่างมากสำหรับการแสวงหาองค์ความรู้ใหม่ในด้านวิทยาศาสตร์ โดยเฉพาะสาขาวิชาคณิตศาสตร์การใช้การอนุมานนิรนัยที่จะช่วยในการพิสูจน์ทฤษฎีบทต่าง ๆ ในขณะที่การอนุมานเชิงประจักษ์ทางวิทยาศาสตร์จะใช้การอนุมานอุปนัยเพื่อการค้นหาความรู้ใหม่

ตัวอย่างที่ 1.3 สมมุติว่าร้านขายเมล็ดพันธุ์เบญจมาศได้มีการเก็บเมล็ดพันธุ์เบญจมาศ 2 ชนิด คือ ชนิดสีขาวและสีเหลือง เพื่อจำหน่ายให้เกษตรกรผู้เพาะปลูกเบญจมาศได้เกิดอุบัติเหตุทำให้เมล็ดเบญจมาศทั้งสองสายพันธุ์ผสมกัน ซึ่งเมล็ดเบญจมาศของสองสายพันธุ์รวมกันเป็น 5 ล้านเมล็ด โดยที่เมล็ดเบญจมาศของสองสายพันธุ์มีลักษณะใกล้เคียงกัน การจะจำแนกว่าเป็นสายพันธุ์ใดได้ จะต้องมีการทดลองปลูกเมล็ดเบญจมาศจนกว่าจะออกดอก จึงจะสามารถจำแนกสายพันธุ์ได้ **จากข้อมูลดังกล่าวหากมีเกษตรกรมาขอซื้อเมล็ดเบญจมาศไปปลูกจากเจ้าของร้านและถามโอกาสที่จะได้เมล็ดเบญจมาศชนิดสีขาวจากเมล็ดทั้งหมด 5 ล้านเมล็ด?**

จากเหตุการณ์นี้จะเห็นได้ว่ามีความเป็นไปได้น้อยมากที่จะทำการทดลองปลูกเมล็ดเบญจมาศทั้งหมด 5 ล้านเมล็ดและรอจนกว่าจะออกดอกเพื่อจำแนกชนิดซึ่งอาจใช้ระยะเวลานาน ดังนั้น แนวทางหนึ่งที่จะได้คำตอบคราว ๆ อาจมีการสุ่มเมล็ดพันธุ์เบญจมาศไปปลูกเพียงบางส่วนแล้วสรุปผลว่าเป็นเบญจมาศชนิดสีขาวอยู่ร้อยละเท่าไร ซึ่งกระบวนการอนุมานนี้ เรียกว่า การอนุมานอุปนัย นั่นคือ คำตอบที่ได้เป็นคำตอบที่ไม่ได้แน่ชัดว่าเป็นเบญจมาศชนิดสีขาวจริง ๆ แต่เป็นเพียงการประมาณค่าภายใต้สถานการณ์ที่ทดลองเท่านั้น ซึ่งมีหลากหลายปัจจัยที่จะส่งผลต่อการสรุปผล เช่น ตัวอย่างสุ่มที่ศึกษามีขนาดที่เหมาะสมและเป็นตัวแทนที่ดีของประชากรที่สนใจศึกษาหรือไม่ แต่การอนุมานอุปนัยก็มีประโยชน์ในหลาย ๆ ด้าน เช่น การประหยัดเวลาและค่าใช้จ่าย เป็นต้น

ดังนั้น ปัญหาในการอนุมานอุปนัยได้ถูกพัฒนาเข้ามาสู่กระบวนการอนุมานเชิงสถิติ โดยเริ่มต้นจากการตั้งวัตถุประสงค์ของการศึกษาจากประชากรเป้าหมายที่สนใจศึกษา ซึ่งอาจทำการศึกษาจากบางส่วนเกี่ยวกับประชากรเป้าหมาย หรือที่เรียกว่า ตัวอย่างสุ่ม เพราะบางครั้งมีข้อกำหนดด้านเวลาและค่าใช้จ่ายในการศึกษาที่ไม่สามารถศึกษาเต็มประชากรได้และบนพื้นฐานของการตรวจสอบบางอย่างอาจมีข้อจำกัดด้านตัวแปรที่สนใจศึกษา เพื่อหาข้อสรุปเกี่ยวกับประชากรเป้าหมายทั้งหมด ซึ่งในบทนี้จะได้ศึกษาเกี่ยวกับการประยุกต์ใช้ความน่าจะเป็นในการอนุมานเชิงสถิติ แนวความคิดพื้นฐานของความน่าจะเป็นและสมบัติที่สำคัญต่อการอนุมานเชิงสถิติ

ตัวอย่าง

1.2 ความน่าจะเป็นและเทคนิคการนับ
(Probability and Counting Techniques)

สถานการณ์ต่าง ๆ ที่แต่ละบุคคลได้ประสบในแต่ละวันมีความไม่แน่นอนเกิดขึ้นตลอดเวลา การพยายามที่จะวัดโอกาสของความไม่แน่นอนของสถานการณ์ต่าง ๆ ที่เกิดขึ้นด้วยตัวเลขโอกาสเกิดขึ้นของเหตุการณ์มีมากน้อยเพียงใด ถ้าตัวเลขโอกาสในการเกิดขึ้นของเหตุการณ์มีค่าเท่ากับ 0 หมายความว่าเหตุการณ์นั้นไม่เกิดขึ้นเลย หรือตัวเลขโอกาสในการเกิดขึ้นของเหตุการณ์เท่ากับ 0.5 หมายความว่าเหตุการณ์นั้นมีโอกาสเกิดขึ้น 50% หรือตัวเลขที่วัดโอกาสในการเกิดขึ้นของเหตุการณ์เท่ากับ 1 หมายความว่าเหตุการณ์นั้นเกิดขึ้นอย่างแน่นอน ซึ่งการวัดโอกาสในการเกิดขึ้นของเหตุการณ์ที่สนใจศึกษามาจากการทดลองที่ไม่ทราบผลลัพธ์ที่จะเกิดขึ้นแน่นอนหรือมีผลลัพธ์ที่เป็นไปได้มากกว่า 1 ผลลัพธ์ เรียกการทดลองนั้นว่า **การทดลองสุ่ม (random experiment)** ซึ่งเรียกเซตหรือกลุ่มของผลลัพธ์ที่เป็นไปได้ทั้งหมดจากการทดลองสุ่มว่า **ปริภูมิตัวอย่าง (sample space)** แทนด้วย S หรือ Ω เมื่อสนใจศึกษา**เหตุการณ์ (event)** คือ สมาชิกหรือกลุ่มของสมาชิกที่สนใจซึ่งเป็นส่วนหนึ่งในปริภูมิตัวอย่าง Ω

ความน่าจะเป็น (probability, พจนานุกรมศัพท์สถิติศาสตร์ ฉบับราชบัณฑิตยสถาน, 2561 หน้า 255) คือ ฟังก์ชันค่าจริง P ที่ส่งจากเซตของเหตุการณ์ที่สนใจศึกษา E ในปริภูมิตัวอย่างที่สนใจศึกษา Ω ไปยังช่วง [0, 1] และมีสมบัติตามสัจพจน์ความน่าจะเป็น ดังนี้

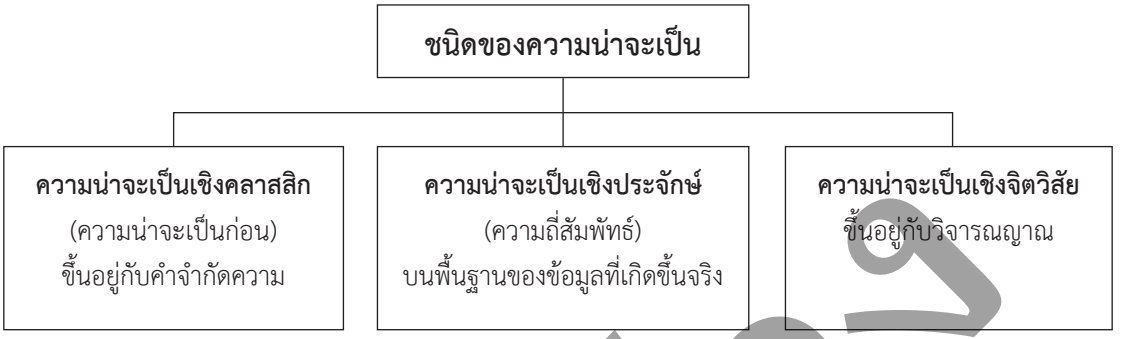
1) $P(E) \geq 0$ ทุกค่า $E \subset \Omega$
2) $P(\Omega) = 1$
3) ถ้า $E_1, E_2, \ldots$ ต่างเป็นเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง Ω และ $E_i \cap E_j = \varnothing$ ทุกค่า $i \neq j$ โดยที่ $i, j = 1, 2, \ldots$ แล้ว

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

ดังนั้น ความน่าจะเป็นของแต่ละเหตุการณ์ใด ๆ ในปริภูมิตัวอย่าง Ω จะมีค่าอยู่ตั้งแต่ 0 ถึง 1 ซึ่งจะเป็นตัวบ่งชี้ว่าเหตุการณ์ที่สนใจศึกษามีโอกาสในการเกิดขึ้นมากน้อยเพียงใด

ตัวอย่าง

สำหรับการคำนวณความน่าจะเป็นของเหตุการณ์ สามารถจำแนกได้เป็น 3 แนวทาง ดังภาพที่ 1.2



ภาพที่ 1.2 ชนิดของความน่าจะเป็น

1) ความน่าจะเป็นเชิงคลาสสิก (classical probability) หรือบางทีเรียกว่าความน่าจะเป็นก่อน (prior probability) คือ ความน่าจะเป็นของเหตุการณ์ที่สนใจศึกษาในปริภูมิตัวอย่าง Ω ของการทดลองสุ่ม โดยที่โอกาสในการเกิดเหตุการณ์ต่าง ๆ จะมีโอกาสเกิดขึ้นเท่า ๆ กัน นั่นคือ การให้น้ำหนักของแต่ละจุดของเหตุการณ์ที่สนใจศึกษามีค่าเท่ากับ $1/n(\Omega)$ เมื่อ $n(\Omega)$ เป็นจำนวนผลลัพธ์ที่เป็นไปได้ทั้งหมด ดังนั้น ความน่าจะเป็นของเหตุการณ์ A ที่สนใจศึกษาในปริภูมิตัวอย่าง Ω เขียนแทนด้วย $P(A)$ ดังนี้

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

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

1) สุ่มหยิบลูกบอล 1 ลูก และได้สีแดง

2) สุ่มหยิบลูกบอล 4 ลูกพร้อมกันและได้สีแดง 2 ลูก

ตัวอย่าง

สถิติศาสตร์ เป็นวิทยาศาสตร์ของการเรียนรู้จากข้อมูล การอนุมานเชิงสถิติอาศัยสารสนเทศที่ได้จากตัวอย่างสุ่ม เพื่อสรุปเกี่ยวกับค่าพารามิเตอร์ที่ไม่ทราบค่าที่ผู้สนใจศึกษา มีข้อสมมุติเกี่ยวกับฟังก์ชัน การแจกแจงความน่าจะเป็นของสิ่งที่สนใจศึกษาไว้ หนังสือเล่มนี้มุ่งเน้น การนำความน่าจะเป็นและสมบัติที่สำคัญ ที่จำเป็นต่อการอนุมานสถิติขั้นสูงต่อไป ไม่ว่าจะเป็นการประมาณค่าพารามิเตอร์แบบจุดและแบบช่วง การตั้งสมมุติฐานและการทดสอบสมมุติฐาน พร้อมทั้งการเสนอแนวทางในการประยุกต์ใช้ความน่าจะเป็น ในการอนุมานเชิงสถิติของผู้เขียนเป็นตัวอย่างประกอบ และการใช้โปรแกรมอาร์ช่วยอธิบายปัญหา สามารถใช้หนังสือเล่มนี้เพื่อประกอบการศึกษาและการค้นคว้าเพิ่มเติมได้

CHIANG MAI UNIVERSITY PRESS

ISBN (e-Book) : 978-616-398-527-9

9 786163 985279

ราคา 400 บาท