

# บทนำและบททบทวน

ชawanai.xinna@exana

ช่างครุตต์สุกศร

ช่างไม้ตัดไม้

บัญทิตฝึกฝนตนเอง

พุทธวจนะ ธรรมนก

## 1.1 ความหมายของสถิติไร้พารามิเตอร์

สถิติไร้พารามิเตอร์ (Nonparametric Statistics) เป็นสถิติอ้างอิง หรือสถิติอนุมานที่ว่า ด้วยการทดสอบสมมติฐาน และการประมาณค่า เช่นเดียวกับสถิติเชิงพารามิเตอร์ (Parametric Statistics) นั่นเอง วิชาสถิติเบื้องต้นโดยทั่วไปจะศึกษาเกี่ยวกับกระบวนการสถิติเชิงพารามิเตอร์ ซึ่งได้แก่แบบทดสอบที่ขึ้นอยู่กับการแจกแจงแบบที (Student t) การวิเคราะห์ความแปรปรวน การวิเคราะห์สหสัมพันธ์ และการวิเคราะห์ทดสอบ โดยเป็นต้น กระบวนการดังกล่าวนี้จะใช้อ้างอิงได้หมาย สมกต่อเมื่อมีข้อตกลงเบื้องต้น (Assumptions) บางอย่างគ្រប់គ្រង ตัวอย่างเช่น กระบวนการอ้างอิงในการวิเคราะห์ความแปรปรวน (ANOVA) จะมีข้อตกลงเบื้องต้นว่าตัวอย่างทั้งหลายต้องเป็นแบบ สุ่มจากประชากรที่มีการแจกแจงแบบปกติ และมีความแปรปรวนเท่ากันในแต่ละกลุ่ม การวัด กระบวนการสถิติเชิงพารามิเตอร์ส่วนมากจะต้องการสเกลการวัดแบบอันตรภาค (Interval Scale) เป็นอย่างน้อย

เนื่องจากประชากรไม่ได้มีข้อตกลงเบื้องต้นตามที่กระบวนการสถิติเชิงพารามิเตอร์ ต้องการเสนอไป เรายังต้องการกระบวนการอ้างอิงที่มีความไว้ใจได้ (Valid) และไม่ขึ้นอยู่กับข้อ ตกลงเบื้องต้นเช่นนั้น กระบวนการดังกล่าวก็คือ กระบวนการสถิติไร้พารามิเตอร์ ซึ่งเป็นกระบวนการ การที่เหมาะสม นឹងจากมีความไว้ใจได้ภายใต้ข้อตกลงเบื้องต้นทั่วไป

กระบวนการสถิติไร้พารามิเตอร์นี้จะประกอบด้วย (1) กระบวนการไร้พารามิเตอร์จริงๆ และ (2) กระบวนการแจกแจงอิสระ กระบวนการไร้พารามิเตอร์ (Nonparametric Procedure) จะไม่ เกี่ยวกับพารามิเตอร์ประชากร นั่นคือไม่มีข้อตกลงเบื้องต้นเกี่ยวกับค่าของพารามิเตอร์ในการ แจกแจงประชากร ตัวอย่างเช่น แบบทดสอบโคสแควร์เกี่ยวกับความเป็นอิสระแบบทดสอบการปรับ ที่ดี (Goodness-of-fit Test) และแบบทดสอบการสุ่ม (Randomness Test) ซึ่งเกี่ยวกับคุณลักษณะ อื่นๆ ที่มิใช่ค่าของพารามิเตอร์ประชากร สำหรับกระบวนการแจกแจงอิสระ (Distribution-Free Procedure) นั้นจะไม่ต้องการข้อตกลงเบื้องต้นใดๆ เกี่ยวกับรูปร่าง หรือการแจกแจงประชากรที่ ตัวอย่างนั้นสุ่มมา ในการประมาณค่าเฉลี่ยประชากร ( $\mu$ ) เราใช้การแจกแจงแบบที่ ถ้าประชากร เป็นแบบปกติ แต่เราสามารถใช้สมการเชิงบีเชฟ (Chebyshev Inequality) ทำการประมาณค่า

เฉลี่ยประชากรได้ ถ้าทราบความแปรปรวนประชากร โดยที่เราไม่จำเป็นต้องทราบรูป่างของประชากรเลย ดังนั้носการเชบบีเชฟจึงเป็นกระบวนการเชิงพารามิเตอร์ แต่เป็นแบบแจกแจงอิสระ โคโนเวอร์ (W.J. Conover) ได้ให้นิยามของวิธีการสถิติไร้พารามิเตอร์ไว้ว่า “วิธีการสถิติ (Statistical Method) จะเป็นแบบไร้พารามิเตอร์ ถ้าสอดคล้องอย่างน้อยเกณฑ์หนึ่งจากเกณฑ์ต่อไปนี้

(1) วิธีการอาจจะใช้กับข้อมูลที่มีสเกลการวัดแบบนามบัญญัติ (2) วิธีการอาจจะใช้กับข้อมูลที่มีสเกลการวัดแบบอันดับ และ (3) วิธีการอาจจะใช้กับข้อมูลที่มีสเกลการวัดแบบอันตรภาคหรืออัตราส่วน โดยที่รูป่างของประชากรจะไม่ระบุ หรือระบุก็ได้ยกเว้นแต่จำนวนพารามิเตอร์ที่ไม่ทราบค่านั้นมีไม่จำกัด” ตามนิยามนี้และโดยทั่วไป สถิติไร้พารามิเตอร์จะเกี่ยวข้องกับสเกลการวัด ดังนั้นจึงจำเป็นต้องกล่าวถึงสเกลการวัดในหัวข้อต่อๆ ไป

## 1.2 ข้อได้เปรียบและข้อเสียเปรียบของสถิติไร้พารามิเตอร์

กระบวนการสถิติไร้พารามิเตอร์มีข้อได้เปรียบและข้อเสียเปรียบ ดังนี้

(1) เนื่องจากกระบวนการไร้พารามิเตอร์นี้อยู่กับข้อตกลงเพียงส่วนน้อย โอกาสที่จะใช้อย่างไม่ปกติจึงน้อย

(2) กระบวนการไร้พารามิเตอร์บางวิธีการนั้น การคำนวณทำได้ง่ายและรวดเร็ว โดยเฉพาะเมื่อทำด้วยมือ ดังนั้นจึงประหยัดเวลาคำนวณ สิ่งที่ควรพิจารณา ก็คือ ถ้าต้องการผลวิเคราะห์อย่างรวดเร็ว หรือไม่มีเครื่องคำนวณชนิดประสิทธิภาพสูง แล้วก็ควรใช้กระบวนการไร้พารามิเตอร์

(3) นักวิจัยที่มีพื้นฐานทางคณิตศาสตร์ และสถิติไม่มาก จะสามารถเข้าใจหลักการและวิธีการของกระบวนการไร้พารามิเตอร์ได้ง่าย

(4) กระบวนการไร้พารามิเตอร์สามารถประยุกต์ใช้กับข้อมูลที่มีสเกลการวัดระดับต่ำได้ เช่นข้อมูลแบบนามบัญญัติหรือแบบอันดับ

(5) ถ้ากระบวนการพารามิเตอร์เหมาะสมมากกว่า แต่เราใช้กระบวนการไร้พารามิเตอร์ เพราะการคำนวณง่ายและรวดเร็ว และจะสูญเสียช่วงสารไป

(6) ถึงแม้กระบวนการไร้พารามิเตอร์จะมีวิธีการคำนวณง่ายๆ แต่มักจะทำซ้ำซากและน่าเบื่อหน่าย

อย่างไรก็ตาม เมื่อต้องการใช้กระบวนการไร้พารามิเตอร์ เรายังมีข้อควรคำนึง ดังนี้

- (1) สมมติฐานที่จะทดสอบนั้นไม่เกี่ยวข้องกับพารามิเตอร์ประชากร
- (2) ข้อมูลที่วัดได้มีสเกลต่ำกว่าที่ต้องการในกระบวนการเชิงพารามิเตอร์ เช่นข้อมูลที่เป็นแบบนับ (นามบัญญัติ) หรือแบบอันดับ
- (3) ข้อตกลงเบื้องต้นที่จำเป็นสำหรับใช้ในกระบวนการเชิงพารามิเตอร์ไม่ครบ
- (4) ผลการวิเคราะห์จะเป็นต้องเชื่อย่างรีบด่วน และการคำนวณต้องทำด้วยมือ

### 1.3 สเกลการวัด (Measurement Scales)

นักสถิติและวิจัยที่วิเคราะห์ข้อมูลชนิดเป็นตัวเลขนั้นจะเกี่ยวข้องกับสเกลที่ใช้ทำการวัด และส่วนมากจะมีทัศนะเกี่ยวกับการวัด และสเกลการวัดตามสตีเวนส์ (S.S. Stevens) สตีเวนส์ได้นิยามการวัด (Measurement) ไว้ว่า “การวัดเป็นการกำหนดตัวเลข (Numerals) ให้แก่เหตุการณ์ หรือวัตถุสิ่งของตามกฎเกณฑ์ (Rule) ที่ระบุไว้” กฎเกณฑ์ที่ระบุหรือตั้งไว้นั้น เราจะเรียกว่า สเกล หรือมาตราการนับวัด (Scale) ของเหตุการณ์นั้น ดังนั้นสเกลจะทำหน้าที่แสดงความสัมพันธ์ระหว่างตัวเลขกับเหตุการณ์ โดยทั่วไปกฎเกณฑ์หรือสเกลการวัดที่ใช้กำหนดตัวเลข จะมีอยู่ 4 แบบ คือ สเกลนามบัญญัติ สเกลเรียงอันดับ สเกลอันตรภาค และสเกลอัตราส่วน ในการกำหนดตัวเลขให้แก่เหตุการณ์นั้นจะใช้กฎเกณฑ์หรือสเกลใดก็พิจารณาได้จากลักษณะที่สำคัญ ดังนี้

- (ก) เหตุการณ์หรือวัตถุสิ่งของนั้นสองอย่างสามารถพิจารณาได้ใหม่ว่าเท่ากันหรือไม่เท่ากัน?
- (ข) เหตุการณ์หรือวัตถุสิ่งของนั้นสองอย่างสามารถให้อันดับที่ได้หรือไม่?
- (ค) สามารถพิจารณาระยะทางระหว่างเหตุการณ์หรือวัตถุสิ่งของได้หรือไม่?
- (ง) สเกลมีจุดที่เป็นศูนย์แท้ (Absolute Zero Point) หรือไม่?

ลักษณะที่สำคัญ	(ก)	(ข)	(ค)	(ง)
สเกลนามบัญญัติ	✓ yes	x no	x	x
เรียงอันดับ	✓	✓	x	x
อันตรภาค	✓	✓	✓	x
อัตราส่วน	✓	✓	✓	✓

สำหรับสเกลการวัด 4 แบบนั้นดังแต่ประเภทหมายเหตุจนถึงละเอียดสุด มีรายละเอียดในแต่ละสเกล ดังนี้

(1) **สเกลนามบัญญัติ (Nominal Scale)** สเกลนี้ใช้ชนาานชื่อให้กับเหตุการณ์หรือวัตถุสิ่งของต่างๆ เช่น เพศก็ชนาานชื่อเป็น 0 (ชาย) และ 1 (หญิง) หรือกลุ่มเลือดก็ชนาานชื่อเป็น 1 (A), 2 (B), 3 (AB), และ 4 (O) เป็นต้น สเกลนามบัญญัตินี้เป็นแบบแรกและดั้งเดิมที่สุด หลักของสเกล แบบนี้ถือแต่เพียงว่า เหตุการณ์ต่างๆ ในสกุลเดียวกันนั้นมีคุณค่าเหมือนกัน และอาจสับเปลี่ยนกัน หรือทดแทนกันได้ เช่น O (หญิง) และ 1 (ชาย) ก็ได้ อย่างไรก็ตาม ถ้าได้กำหนดตัวเลขใดหมายถึงเหตุการณ์อะไรแล้วก็อย่าใช้ตัวเลขนั้นไปช้ากับเหตุการณ์อื่นๆ เช้าอีกก็แล้วกันและโดยทั่วไปเดียวกัน ในเหตุการณ์หนึ่งอย่าใช้ตัวเลขเกินกว่า 1 จำนวนไปแทนเหตุการณ์นั้นเช้าอีก เพราะจะทำให้eng ไม่รู้ว่าตัวเลขตัวไหนแทนเหตุการณ์ใดกันแน่

(2) **สเกลเรียงอันดับ (Ordinal Scale)** สเกลนี้ใช้เรียงหรือจัดอันดับเหตุการณ์ต่างๆ ในสกุลเดียวกัน ให้ลดหลั่นกันเป็นชั้นๆ ตามปริมาณและคุณภาพมากน้อยอีกด้วย ดังเช่นเมื่อตั้งนามบัญญัติให้มนุษย์เป็น ชาย (0) และ หญิง (1) แล้วก็ยังชอยหรือแจงประเภทของหญิงออกเป็น

อันดับฯ ว่า 1 (สวยมาก), 2 (ปานกลาง) และ 3 (ชี้หรรษ) หรือแบ่งช้ายออกเป็น 1 (อ่อน), 2 (สมทรง) และ 3 (ผอม) หรือแบ่งนักเรียนในชั้นเดียวกันออกเป็น 1 (ใจ), 2 (ปานกลาง) และ 3 (ฉลาด) เป็นต้น อย่างไรก็ตาม สเกลชนิดนี้เป็นเพียงการจัดอันดับให้ลดหลั่นกันเป็นชั้นๆ เท่านั้นยังไม่สามารถบอกได้ว่า แต่ละชั้นห่างกันเท่าไร และทุกๆ ชั้นห่างกันหรือไม่

(3) **สเกลอันตรภาค (Interval Scale)** สเกลนี้ใช้แบ่งหรือกำหนดเหตุการณ์หรือวัตถุสิ่งของในสกุลเดียวกันออกเป็นประเภทหรือเป็นจังหวะ เป็นช่วงๆ ชั้นๆ ที่มีขนาดใหญ่เท่าๆ กัน เช่นแบ่งเวลา 1 ปี ออกเป็น 365 วัน หรือ 365 ช่วงที่มีความนาฬาเท่าๆ กัน และแบ่งก้านเทอร์โม-มิเตอร์ ชนิดเซลเซียสออกเป็น 100 องศา หรือ 100 ช่องที่เท่าๆ กัน เป็นต้น อย่างไรก็ตาม ตัวเลขที่กำหนดนั้นถ้าเป็น 0 ก็ไม่ได้หมายความเป็นศูนย์แท้หรือศูนย์อนันต์ (True or Absolute Zero) แต่เป็นศูนย์เทียมหรือศูนย์สัมพัทธ์ (Arbitrary or Relative Zero) เท่านั้น และจำนวนเท่าจะไม่เป็นจริง เช่นคะแนนสอบถ้าได้ 0 ก็ไม่ได้หมายความว่าไม่มีความรู้ หรือได้ 50 ก็ไม่ได้หมายความว่ามีความรู้เป็น 2 เท่าของคะแนน 25 เป็นต้น

(4) **สเกลอัตราส่วน (Ratio Scale)** สเกลนี้สมบูรณ์ที่สุด ทางฟิสิกส์สามารถสร้างสเกลนี้ได้ ดังเช่นสเกลความยาว น้ำหนัก ความหนาแน่น และความด้านทานไฟฟ้า เป็นต้น คุณสมบัติพิเศษของสเกลนี้คือจุดศูนย์แท้ หรือศูนย์อนันต์ และมีหน่วยเป็นขนาดโดยเท่าๆ กัน เรียงซึ่งลงตามลำดับสม่ำเสมอ กัน และจำนวนเท่าของค่าจะเป็นจริง เช่นน้ำหนัก 0 กิโลกรัมไม่มีน้ำหนักอะไรเลย และน้ำหนัก 2, 3, 4 กิโลกรัม กิโลหนักเป็น 2, 3 และ 4 เท่าตามลำดับของ 1 กิโลกรัม แห่งๆ ด้วย เป็นต้น

#### 1.4 การทดสอบสมมติฐาน (Hypothesis Testing)

การทดสอบสมมติฐานเป็นเรื่องหนึ่งของสถิติอนุมาน ซึ่งในสถิติไว้พารามิเตอร์ ส่วนใหญ่จะศึกษาเกี่ยวกับการทดสอบสมมติฐาน คำว่า สมมติฐาน (Hypothesis) นั้นหมายถึงคำกล่าวเที่ยวกับประชากรหนึ่งประชากรหรือมากกว่า โดยทั่วไปเรามักจะกล่าวถึงสมมติฐานอยู่ 2 แบบ คือ สมมติฐานวิจัย (Research Hypothesis) และ สมมติฐานสถิติ (Statistical Hypothesis) สมมติฐานวิจัยนั้นกำหนดหรือสร้างขึ้นจากนักสำรวจ นักสืบสวน หรือนักทดลอง ซึ่งโดยปกติมิใช่นักสถิติ ป้อยครั้งที่สมมติฐานวิจัยเป็นจากการสังหารณ์ หรือความสงสัย ซึ่งขึ้นอยู่กับการสังเกตอย่างกว้างขวางของนักสืบสวน เช่นครูที่สอนมานานอาจจะสงสัยว่า สภาพภารณ์บางอย่างในห้องเรียนขาดช่วงต่อการเรียน เป็นต้น ความสงสัยนี้จะนำไปสู่ สมมติฐานวิจัย เช่น “นักเรียนจะได้คะแนนสูงในวิชาเลขคณิตถ้าอุณหภูมิของห้องเรียนไม่เกิน 68° F” เป็นต้น

สำหรับ สมมติฐานสถิติ (Statistical Hypothesis) นั้นประกอบด้วย สมมติฐานหลัก (Null Hypothesis) ซึ่งแทนด้วย  $H_0$  และ สมมติฐานรอง (Alternative Hypothesis) ซึ่งแทนด้วย  $H_a$

สมมติฐานหลักจะเป็นสมมติฐานที่เราจะทดสอบนั่นเอง กระบวนการทดสอบ (Test Procedure) ซึ่งขึ้นอยู่กับข่าวสารที่พัฒนาจากข้อมูลของตัวอย่างที่เหมาะสมนั้นจะให้ผลในการตัดสินใจประการหนึ่งในสองประการคือ (1) ตัดสินใจปฏิเสธสมมติฐานหลัก  $H_0$  หรือ (2) ตัดสินใจไม่ปฏิเสธ สมมติฐานหลัก เพราะว่าตัวอย่างไม่บรรจุประจักษ์พยานเพียงพอที่จะประกันการปฏิเสธ เมื่อปฏิเสธ สมมติฐานหลัก เรายังยอมรับสมมติฐานรองว่าเป็นจริง เราสามารถทำเช่นนี้ได้ เพราะว่าหากล่าว สมมติฐานหลัก และสมมติฐานรองให้แยกกัน แต่เป็นส่วนประกอบ (Mutually exclusive and Complementary) โดยปกติแท้ไม่เสมอไป เราจะเห็นว่าสมมติฐานรองและสมมติฐานวิจัยเป็นแบบเดียวกัน ดังนั้นส่วนมากสมมติฐานรองจะเป็นคำกล่าวของสิ่งหรือเหตุการณ์ที่เราคาดหวังว่าสามารถสรุปได้

สมมติฐานที่ทดสอบอาจจะเป็นสองทาง (Two-sided) ซึ่งไม่มีทิศทาง (Nondirectional) หรือทางเดียว (One-sided) ซึ่งมีทิศทาง (Directional) การทดสอบสมมติฐานนั้นผู้สืบสวนจะเลือกตัวสถิติทดสอบ (Test Statistic) ที่เหมาะสม และระบุการแจกแจงของมัน เมื่อ  $H_0$  เป็นจริง จากข้อมูลตัวอย่างที่สังเกตได้เราคำนวณค่าของตัวสถิติทดสอบ และพิจารณาว่า “เป็นค่าที่สุดเหวี่ยง (Surprising Extreme) ที่สังเกตได้เมื่อ  $H_0$  เป็นจริงหรือไม่ (ค่าสุดเหวี่ยงนั้นก็คือค่าที่มากเกินไปหรือน้อยเกินไปนั่นเอง)” หรือว่า “ขนาดของค่าที่คำนวณได้ของตัวสถิติทดสอบนั้นสุดเหวี่ยงเพียงพอที่จะทำให้ปฏิเสธสมมติฐานหลักหรือไม่” ก่อนที่จะตรวจสอบข้อมูลตัวอย่างผู้สืบสวนส่วนมากจะตั้งกฎตัดสินใจไว้ กฎนี้จะกล่าวว่า “จะปฏิเสธ  $H_0$  ถ้าความน่าจะเป็นที่จะได้ค่าของตัวสถิติทดสอบมีขนาดตามที่กำหนดไว้ หรือมากกว่า ในเมื่อ  $H_0$  เป็นจริง นั้นเท่ากับหรือน้อยกว่าจำนวนน้อย ๆ  $\alpha$ ” จำนวน  $\alpha$  นั้นได้ชื่อว่า ระดับนัยสำคัญ (Level of Significance) หรือขนาดทดสอบ (Size of the Test) เมื่อใช้วิธีการตัดสินใจก็จะเลือก  $\alpha$  เท่ากับ 0.05 หรือ 0.01 หรือบางครั้งเป็น 0.10

ค่าวิกฤต (Critical Value) ของตัวสถิติทดสอบ เป็นค่าที่สุดเหวี่ยง ซึ่งมีโอกาสที่จะได้ค่านั้น หรือได้ค่าสุดเหวี่ยงกว่า เป็นโอกาสเท่ากับ  $\alpha$  เมื่อ  $H_0$  เป็นจริง ดังนั้นความสามารถล่ากว่ากฎตัดสินใจในเทอมของค่าวิกฤตได้ ตัวอย่างเช่น สำหรับการทดสอบทางเดียว กฎตัดสินใจจะแนะนำให้เราปฏิเสธ  $H_0$  ถ้าค่าที่คำนวณได้ของตัวสถิติทดสอบนั้นสุดเหวี่ยง (มากกว่าหรือน้อยกว่า แล้วแต่ทิศทางของสมมติฐานรอง) กว่าค่าวิกฤต สำหรับการทดสอบสองทางจะมีสองค่าวิกฤต คือเมื่อเราปฏิเสธ  $H_0$  ถ้าค่าที่คำนวณได้ของตัวสถิติทดสอบมากกว่าค่าที่มาก ของค่าวิกฤตสองค่าหรือน้อยกว่าค่าที่น้อยของสองค่าวิกฤต

**1.4.1 ค่า P (P Values)** มีวิธีการที่จะตัดสินใจว่าข้อมูลตัวอย่างนั้นทำให้เกิดความสังสัยในสมมติฐานหลักหรือไม่ นั่นคือการพิจารณาความน่าจะเป็นที่ได้ค่าของตัวสถิติทดสอบนั้นอย่างน้อยสุดเหวี่ยงเท่ากับค่าที่สังเกตได้จริง ในเมื่อ  $H_0$  เป็นจริง ถ้าให้  $G$  เป็นตัวสถิติทดสอบ และ  $G_u$  เป็นค่าที่คำนวณได้ของ  $G$  จากข้อมูลตัวอย่าง แล้วค่า  $P$  จะเป็นอย่างหนึ่งอย่างใด ตามสมมติฐานรองดังนี้

$$p = P(G \leq G_c), p = P(G \geq G_c),$$

หรือ  $p = 2\min \{ P(G \leq G_c), P(G \geq G_c) \}$

ความน่าจะเป็นหรือค่า  $P$  นี้มีชื่อเรียกหลายชื่อคือ Critical Value, The Descriptive Level of Significance, The Associated Probability, The Prob Value, The Significance Probability ยอดจ์และเลห์แมน (J.L. Hodges., E.L. Lehman) ได้แนะนำว่า ค่า  $P$  เป็นจำนวนเลขที่เป็นมาตรฐานวัดตีกรีของความประหลาดใจที่การทดลองจะทำให้ผู้เชื่อถือในสมมติฐานหลัก สำหรับค่า  $P$  นั้นจะรายงานได้ตั้งเทอมต่อไปนี้  $p > 0.05$ ,  $p < 0.01$ ,  $0.01 < p < 0.05$  หรือ  $p = 0.065$  ในเมื่อ  $p$  เป็นค่า  $P$  ดังนั้นค่า  $P$  สามารถรายงานได้ในรูปของค่าแท้จริง หรือเป็นช่วง ทั้งนี้แล้วแต่ตารางที่ให้ไว้ของ การแจกแจงของตัวสถิติทดสอบ

ตัวอย่าง (1) สมมติว่าสมมติฐานที่จะทดสอบเป็น

$$H_0: \mu \leq \mu_0; H_a: \mu > \mu_0$$

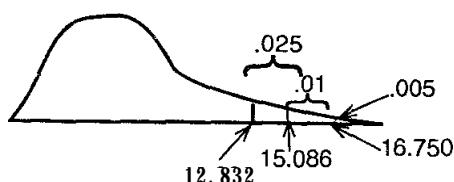
ตัวสถิติทดสอบเป็น  $Z$  ซึ่งมีการแจกแจงแบบปกติมาตรฐาน และค่าที่คำนวณได้ของ  $Z$  เป็น 1.86 แล้วเราจะได้ค่า  $P$  เป็น

$$\begin{aligned} p &= P(Z \geq 1.86, H_0 \text{ เป็นจริง}) \\ &= 0.5 - 0.4686 = 0.0314 \end{aligned}$$

(2) สมมติว่าในการทดลองนั้น ตัวสถิติทดสอบแจกแจงแบบโคสแคร์ ด้วยองค์ความ เป็นอิสระ 5 และค่าที่คำนวณได้ของตัวสถิติทดสอบเป็น 14.665 จากตารางโคสแคร์ เราพบว่า 14.665 อยู่ระหว่าง  $\chi_{0.025}^{2(5)} = 12.832$  และ  $\chi_{0.01}^{2(5)} = 15.086$  ดังนั้นเราจึงรายงานค่า  $P$  เป็น  $0.01 < p < 0.025$

แต่ถ้าได้ค่าของตัวสถิติทดสอบเป็น 17.335 ซึ่งค่านี้มากกว่า  $\chi_{0.005}^{2(5)} = 16.750$  แล้วเรา เชียนค่า  $P$  ได้เป็น  $p < 0.005$

สำหรับค่า  $p$  จากทั้งสองตัวอย่างนั้นแสดงได้ดังนี้



สำหรับค่า  $p$  นั้นจะให้ข่าวสารมากกว่าคำกล่าวที่ว่า “ความแตกต่างมีนัยสำคัญ ณ ระดับ 0.05” หรือ “ $H_0$  สามารถปฏิเสธได้ ณ ระดับ 0.01” อันนี้ก็เป็นหลักสำคัญสำหรับการรายงานค่า  $p$  ในงานวิจัย เพื่อกำหนดค่า  $p$  ให้ เราสามารถเลือกระดับนัยสำคัญของเราได้เอง หรือเลือกระดับที่ เรายึดมั่นที่จะเลิกเชื่อว่า  $H_0$  เป็นจริง และเริ่มเชื่อว่า  $H_a$  เป็นจริง

**1.4.2 อำนาจทดสอบ และประสิทธิภาพของแบบทดสอบสมมติฐาน (Power and Efficiency of a Hypothesis Test)** เกณฑ์ในการประเมินผลของแบบทดสอบที่เรารู้จัก กันดีก็คืออำนาจของแบบทดสอบสมมติฐานซึ่งเป็นความน่าจะเป็นที่จะปฏิเสธสมมติฐานหลักที่เป็นเท็จ อำนาจทดสอบนั้นอาจจะกำหนดได้เป็น  $1 - \beta$  ในเมื่อ  $\beta$  เป็นความน่าจะเป็นที่จะยอมรับสมมติฐานหลักที่เป็นเท็จ การยอมรับสมมติฐานหลักที่เป็นเท็จนั้นได้ซึ่งว่า ความคลาดเคลื่อนประเททสอง (Type II Error) และการปฏิเสธสมมติฐานหลักที่เป็นจริง จะได้ซึ่งว่า ความคลาดเคลื่อนประเททหนึ่ง (Type I Error) ความน่าจะเป็นของความคลาดเคลื่อนประเททหนึ่งจะแทนด้วย  $\alpha$  โดยทั่วไปอำนาจทดสอบสูงเป็นคุณลักษณะที่ต้องการของแบบทดสอบ

เกณฑ์ในการประเมินแบบทดสอบอีกเกณฑ์หนึ่งก็คือ ประสิทธิภาพ (Efficiency) ตัวนี้ ของประสิทธิภาพในแบบทดสอบไร้พารามิเตอร์ก็คือประสิทธิภาพสัมพัทธ์นิดตัวอย่างขนาดโต (ARE' Asymptotic Relative Efficiency) หรือ ประสิทธิภาพแบบปีทแมน (Pitman Efficiency) คุณสมบัติที่พึงประสงค์ของแบบทดสอบก็คือ มีประสิทธิภาพสูง โดยทั่วไป ARE ของแบบทดสอบ เป็นการประมาณค่าของประสิทธิภาพสัมพัทธ์ ประสิทธิภาพสัมพัทธ์ของแบบทดสอบ A ต่อแบบทดสอบ B (สำหรับ  $H_0, H_A \alpha$  และ  $\beta$  เดียวกัน) จะเป็นอัตราส่วน  $g_A/g_B$  ในเมื่อ  $g_A$  และ  $g_B$  เป็น ขนาดตัวอย่างของแบบทดสอบ A และ B ถ้า  $g_A$  น้อยกว่า  $g_B$  แล้วประสิทธิภาพของแบบทดสอบ A สัมพันธ์กับแบบทดสอบ B จะมากกว่าหนึ่ง ซึ่งเราจะพูดว่าแบบทดสอบ A มีประสิทธิภาพกว่าแบบทดสอบ B แบบทดสอบที่ต้องการขนาดตัวอย่างที่น้อยกว่าภายในได้เงินใช้เดียวกันจะดีกว่า เพราะ ตัวอย่างน้อยจะลดค่าใช้จ่าย เวลา และทรัพยากรอื่นๆ

## 1.5 การประมาณค่า (Estimation)

บ่อยครั้งเราต้องการตัดสินใจเกี่ยวกับค่าของพารามิเตอร์ประชากรกว่าที่จะทราบว่า จะปฏิเสธสมมติฐานหลักที่ว่าค่านั้นเท่ากับค่าที่ระบุไว้หรือไม่ การตัดสินใจเกี่ยวกับขนาดของพารามิเตอร์ประชากรที่ชื่นอยู่กับข้อมูลตัวอย่างนั้น เราจะใช้กระบวนการของการประมาณค่า

การประมาณค่ามีอยู่ 2 แบบ คือการประมาณค่าแบบจุด (Point Estimation) และการประมาณค่าแบบช่วง (Interval Estimation) ในการประมาณค่าแบบจุด เราจะคำนวณค่าเดียวๆ ที่ เรียกว่า ค่าประมาณ (Estimate) จากข้อมูลตัวอย่าง และใช้ค่านี้เป็นตัวแทน (Candidate) สำหรับ พารามิเตอร์ที่เราต้องการประมาณ ส่วนมากค่าประมาณแบบช่วงเป็นค่าประมาณที่ต้องการและมี ประโยชน์มากกว่า ค่าประมาณแบบช่วงประกอบด้วยค่าที่เป็นไปได้ 2 ค่าของพารามิเตอร์ที่ต้องการ ประมาณ (นั่นคือค่าต่ำและค่าสูง) สองค่านี้จะกำหนดช่วงที่ทำให้เราแสดงถึงตัวของความเชื่อมั่นที่ ว่า ช่วงนั้นจะบรรจุพารามิเตอร์ที่ประมาณไว้ด้วยได้ เพราะฉะนั้นค่าประมาณแบบช่วงจึงเรียกวัน บอยๆ ว่า ช่วงเชื่อมั่น (Confidence Interval)

เราแสดงตีกริของความเชื่อมั่นในช่วงเชื่อมั่นโดยใช้สัมประสิทธิ์เชื่อมั่น (Confidence Coefficient) ซึ่งเป็นจำนวนเลขอยู่ระหว่าง 0 และ 1 หรือเป็นเปอร์เซนต์ ตัวอย่างเช่น ถ้าสัมประสิทธิ์เชื่อมั่นเป็น 0.95 หรือ 95% เราจะพูดว่าเราเชื่อมั่น 95% ที่ช่วงนั้นจะบรรจุพารามิเตอร์ที่เรากำลังประมาณ

เราสร้างช่วงซึ่งสามารถแปลความหมายได้ 2 วิธี คือการแปลความหมายเชิงความน่าจะเป็น (Probabilistic Interpretation) ซึ่งขึ้นอยู่กับความจริงที่ว่า ใน การสุ่ม抽 100 (1- $\alpha$ )% ของช่วงที่สร้างขึ้นแบบเดียวกัน (และตัวอย่างขนาดเดียวกัน) จะบรรจุพารามิเตอร์ที่ประมาณ การแปลความหมายเช่นนี้ใช้กับช่วงเชื่อมั่นทั้งหมดที่เราสามารถสร้างได้ ในทางปฏิบัติเราสร้าง เพียงช่วงเดียว และเป็นช่วงเดียวที่เราใช้แปลความหมายแบบอื่นนั้นคือการแปลความหมายเชิงปฏิบัติ (Practical Interpretation) ในการแสดงความหมายเชิงปฏิบัตินี้เราเชื่อมั่น 100 (1- $\alpha$ )% ว่าช่วงเดียวที่สร้างขึ้นจะบรรจุพารามิเตอร์ที่เราประมาณ การแปลความหมายทั้งสองนี้ 100 (1- $\alpha$ )% เป็นสัมประสิทธิ์ เชื่อมั่น

เราสามารถแสดงช่วงเชื่อมั่นสำหรับพารามิเตอร์ ในเทอมความน่าจะเป็น ได้เป็น

$$P(L_0 < \theta < U_0) = 1 - \alpha$$

ในเมื่อ  $L_0$  และ  $U_0$  เป็นตัวแปรเชิงสุ่มที่สอดคล้องกับคำกล่าวเชิงน่าจะเป็น เมื่อเราระบุค่าของ  $L_0$  และ  $U_0$  เป็น  $L$  และ  $U$  แล้วช่วงนั้นก็ไม่เป็นตัวแปรเชิงสุ่มอีกต่อไป แต่จะเป็นช่วงที่จำกัดนั้นคือเป็นช่วงเดียวๆ ที่สร้างขึ้นในการประยุกต์เชิงปฏิบัติ เราสามารถแสดงช่วงเดียวๆ อย่างกะทัดรัดได้เป็น

$$C(L < \theta < U) = 1 - \alpha$$

ในเมื่อ  $C$  แทนความเชื่อมั่น และช่วงคำกล่าววนนี้เป็นคำกล่าวเชิงเชื่อมั่น ไม่ใช่คำกล่าวเชิงน่าจะเป็น การประมาณค่าแบบช่วง และการทดสอบสมมติฐานจะสัมพันธ์กัน ซึ่งพิจารณาได้จาก สมมติฐาน

$$H_0 : \theta = \theta_0 ; H_a : \theta \neq \theta_0$$

ในเมื่อระดับนัยสำคัญเป็น  $\alpha$  ค่าที่เป็นไปได้ของพารามิเตอร์  $\theta$  ที่บรรจุในช่วง  $L < \theta < U$  ซึ่งเป็น ช่วงเชื่อมั่น 100 (1- $\alpha$ )% นั้นจะเป็นค่าที่ตรงกันกับสมมติฐานหลัก ค่าที่เป็นไปได้ของ  $\theta$  ที่อยู่นอกช่วงจะไม่ตรงกันกับสมมติฐานหลัก เพราะฉะนั้นเราทดสอบ  $H_0$  โดยอาศัยช่วงเชื่อมั่นได้ ถ้า ช่วงเชื่อมั่น 100 (1- $\alpha$ )% ไม่บรรจุค่าตามสมมติฐานของพารามิเตอร์  $\theta_0$  เราก็ปฏิเสธ  $H_0$  ณ ระดับ นัยสำคัญ  $\alpha$  ถ้า  $\theta_0$  อยู่ในช่วง เราไม่ปฏิเสธ  $H_0$  (ณ ระดับนัยสำคัญ  $\alpha$ )

The University brings out all Abilities  
including Incapability.

*Anton Chekhov*

Education makes a people easy to lead  
but difficult to drive; easy to govern  
but impossible to ensolve.

*Lord Brougham*

Being afraid at the right time is called  
Education.

*Angelo*

## ตัวอักษรกรีก

ตัวอักษรเล็ก	ตัวอักษรใหญ่		ชื่อ
$\alpha$	A	alpha	แอลฟ่า
$\beta$	B	beta	บีتا
$\gamma$	Γ	gamma	แกมมา
$\delta, \partial$	Δ	delta	เดลต้า
$\varepsilon$	Ε	epsilon	เอปไซลอน
$\zeta$	Z	zeta	ซีตา
$\theta$	H	eta	อีตา
$\theta$	Θ	theta	ทีตา
$\iota$	I	iota	ไอโอดา
$\kappa$	K	kappa	แคปปา
$\lambda$	Λ	lambda	แลมบ์ด้า
$\mu$	M	mu	มิว
$\nu$	N	nu	นิว
$\xi$	Ξ	xi	ไซ
$\omicron$	O	omicron	โอมีครอน
$\pi$	Π	pi	พาย
$\rho$	R	rho	โร
$\sigma$	Σ	sigma	ซิกมา
$\tau$	T	tau	เทา
$\upsilon$	Υ	upsilon	อิปไซลอน
$\phi$	Φ	phi	ฟาย
$\chi$	X	chi	ไค
$\psi$	Ψ	psi	ไซ
$\omega$	Ω	omega	โอมega