การจัดการกับข้อมูล
การทดสอบความชำนาญด้วยวิธีทางสถิติ

ศิริวรรณ ศิลป์สกุลสุข
กานดา โกมลวัฒนชัย

การทดสอบความชำนาญเป็นการใช้เทคนิคการเปรียบเทียบผลการทดสอบระหว่างห้องปฏิบัติการเพื่อหาสมรรถนะของแต่ละห้องปฏิบัติการสำหรับการวิเคราะห์ทดสอบที่กำหนด และยังใช้ในการเฝ้าดูสมรรถนะของห้องปฏิบัติการอย่างต่อเนื่อง โดยขั้นตอนการดำเนินการของการเปรียบเทียบจะต้องใช้ตัวอย่างเดียวกันหรือตัวอย่างที่มีความเป็นเนื้อเดียวกันและมีความเสถียรในการประเมินความสามารถของการวิเคราะห์ทดสอบของแต่ละห้องปฏิบัติการ ดังนั้นการทดสอบความชำนาญจึงเป็นเครื่องมือที่ใช้ในการควบคุมคุณภาพของการวิเคราะห์ทดสอบจากภายนอกได้เป็นอย่างดี และห้องปฏิบัติการที่เข้าร่วมกิจกรรมทดสอบความชำนาญจะได้รับความน่าเชื่อถือและยอมรับซึ่งกันและกัน

ผลจากการทดสอบความชำนาญสามารถแสดงได้หลายรูปแบบขึ้นกับลักษณะและการกระจายเชิงสถิติของข้อมูล การเลือกใช้เทคนิคทางสถิติแบบใดในการวิเคราะห์ผลการทดสอบความชำนาญขึ้นกับความเหมาะสมของสถานการณ์และไม่อาจระบุวิธีใดเป็นการเฉพาะ อย่างไรก็ตามเมื่อผลของผู้เข้าร่วมโครงการถูกนำมาประเมินจะ มี 3 ขั้นตอนพื้นฐานที่สามารถใช้ได้กับทุกการทดสอบความชำนาญ[3] คือ

1. การกำหนดค่า (assigned value)
2. การคำนวณความสามารถเชิงสถิติ (Performance statistic)
3. การประเมินความสามารถ (Evaluation of Performance)

และในบางกรณีอาจจะต้องทำการทดสอบเบื้องต้นว่าตัวอย่างที่ใช้ทดสอบมีความเป็นเนื้อเดียวกันและมีความเสถียร

สำหรับโครงการการเปรียบเทียบความสามารถห้องปฏิบัติการ (Interlaboratory comparison scheme) ที่เป็นโครงการใหม่ ผลการประเมินในระยะแรกๆ มักจะไม่ค่อยดี ทั้งนี้อาจจะเกิดจากแบบฟอร์มที่เพิ่งเริ่มใช้ ตัวอย่างที่ใช้ทดสอบ วิธีทดสอบ หรือขั้นตอนที่แตกต่างกันในแต่ละห้องปฏิบัติการ ผู้จัดทำโครงการต้องใช้ วิธี robust ในการเปรียบเทียบสมรรถนะ (เช่น เปอร์เซ็นต์ไตล์) จนกระทั่งผลการประเมินเริ่มดีขึ้น เมื่อผลการประเมินระหว่างห้องปฏิบัติการเริ่มดีขึ้น โครงการการทดสอบความชำนาญจะต้องมีการกำหนดขั้นตอนอย่างชัดเจนและต้องมีการระบุเทคนิคทางสถิติที่ใช้

การกำหนดค่า Assigned Value

ค่า assigned values คือค่าที่กำหนดขึ้นเพื่อใช้ในการประเมิน ผู้เข้าร่วมโครงการอย่างยุติธรรม และยังเป็นการสนับสนุนความสอดคล้องกันระหว่างห้องปฏิบัติการต่างๆ และระหว่างวิธีการต่างๆ ที่ใช้ซึ่งสามารถทำได้โดยการเลือกกลุ่มที่จะทำการเปรียบเทียบกันให้อยู่ในระดับความสามารถเดียวกัน และ นำค่า assigned values มาใช้เป็นค่ากลาง

วิธีการที่ใช้ในการกำหนดค่า assigned value มีหลายวิธี ส่วนหนึ่งของวิธีการที่นิยมใช้กันมากได้แก่

1. Known values เป็นค่าที่ได้จากผลการทดสอบวัตถุที่จัดเตรียมขึ้นโดยเฉพาะ (เช่น ผลิตจากโรงงาน หรือจากการเจือจางสารที่ทราบความเข้มข้นที่แน่นอน)

2. Certified reference values เป็นค่าที่ได้จากวิธีทดสอบที่เฉพาะเจาะจง (สำหรับการทดสอบเชิงปริมาณ)

3. Reference values เป็นค่าที่ได้จากการวิเคราะห์ การวัดปริมาณ หรือจากผลการเปรียบเทียบของวัตถุทดสอบกับวัตถุอ้างอิง (หรือวัสดุมาตรฐาน) ที่สามารถทวนสอบได้ถึงมาตรฐานระดับชาติหรือระดับสากล

4. Consensus values เป็นค่าที่ได้จากห้องปฏิบัติการที่มีความชำนาญ โดยห้องปฏิบัติการเหล่านั้นต้องแสดงให้เห็นถึงความสามารถในการทดสอบตามรายการทดสอบที่ระบุ โดยใช้วิธีทดสอบที่ได้ตรวจสอบความใช้ได้แล้ว (validation methods) ว่ามีความแม่นยำและความเที่ยงตรงสูง และสามารถเปรียบเทียบได้กับวิธีที่ทดสอบที่ใช้กันทั่วไป สำหรับห้องปฏิบัติการที่กำหนดให้ทดสอบนั้นบางครั้งอาจจะเป็นห้องปฏิบัติการอ้างอิง (Reference Laboratories)

5. Consensus values ที่ได้จากห้องปฏิบัติการที่เข้าร่วมโครงการทดสอบความชำนาญ โดยใช้ค่า “เฉลี่ย” ที่ได้จากการเปรียบเทียบภายในกลุ่มที่เหมาะสม ซึ่งค่าเฉลี่ยในทางสถิติสามารถแสดงได้หลายรูปแบบ เช่น

5.1 ค่าเฉลี่ยเลขคณิต หรือมัชฉิมเลขคณิต (Arithmetic mean)

5.2 ค่าเฉลี่ยเรขาคณิต (Geometric mean)

5.3 ค่ามัธยฐาน (median)

5.4 ค่าฐานนิยม (mode)

หรือ การวัด robust แบบอื่น

เมื่อใช้ผลการทดสอบของผู้เข้าร่วมโครงการกำหนดเป็นค่า assigned values จะต้องคำนึงผลของค่าต่างสุด (extreme results) จากค่าอื่นๆ ในชุดข้อมูล ซึ่งสามารถจัดการโดยใช้เทคนิคที่จะลดอิทธิพลจากค่าต่างสุดโดยใช้วิธีทางสถิติแบบโรบัส (robust statistic methods) หรือโดยการตัด outliers (ข้อมูลที่มีค่าผิดปกติอย่างมีนัยสำคัญ) ออกก่อนที่จะทำการคำนวณ ถ้าผลการทดสอบบางค่าถูกตัดออกเป็น outliers ผลการทดสอบค่านั้นจะถูกตัดเพื่อการคำนวณผลสรุปทางสถิติเท่านั้น แต่ยังคงถูกนำมาประเมินในขั้นตอนของการเปรียบเทียบการทดสอบความชำนาญ และจัดระดับความสามารถอยู่

การคำนวณความสามารถเชิงสถิติ (Performance statistic)

ผลจากการทดสอบความชำนาญบ่อยครั้งจะถูกเปลี่ยนให้อยู่ในรูปความสามารถเชิงสถิติ เพื่อใช้ในการแปลผลและเพื่อใช้เปรียบเทียบกับเป้าหมายที่วางไว้ โดยมีวัตถุประสงค์ที่จะวัดความเบี่ยงเบนไปจากค่ากลางที่กำหนดขึ้น (assigned value) และสามารถใช้เป็นเกณฑ์วัดความสามารถ เทคนิคที่ใช้อาจจะเริ่มจากขั้นตอนที่ง่ายๆ ไปจนถึงการใช้สถิติที่มีความซับซ้อน

การวัดการกระจายนิยมใช้เพื่อคำนวณความสามารถเชิงสถิติและใช้ในรายงานสรุปของโครงการทดสอบความชำนาญ ตัวอย่างของการวัดการกระจายสำหรับกลุ่มเปรียบเทียบที่เหมาะสมที่ใช้กันทั่วไป ได้แก่

1. ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation, SD)

2. สัมประสิทธิ์ความแปรผัน (Coefficient of variation, CV) หรือ ส่วนเบี่ยงเบนมาตรฐานสัมพัทธ์ (Relative standard deviation, RSD)

3. เปอร์เซ็นต์ไทล์ (Percentiles)

4. Median Absolute Deviation (MAD) หรือการวัดแบบ robust อื่นๆ

สถิติที่ใช้ในการประเมินความสามารถ โดยลำดับจากขั้นตอนที่ง่ายไปจนถึงการใช้สถิติที่มีความซับซ้อนคือ

1. ผลต่างระหว่าง (x-X) เมื่อ x คือผลการทดสอบของผู้เข้าร่วมโครงการ และ X คือ ค่าของ assigned value

ความแตกต่างระหว่างผลการทดสอบของผู้เข้าร่วมโครงการกับค่า assigned value บางครั้งก็อาจจะเพียงพอในการประเมินความสามารถ และเป็นสิ่งที่ทำให้ผู้เข้าร่วมโครงการเข้าใจได้ง่ายที่สุด ค่า

(x-X) เรียกว่า ค่า “Estimate of laboratory bias”

2. ร้อยละความแตกต่าง

$$\frac{(x-X)}{X} \times 100$$

เป็นผลต่างที่ปรับให้อยู่ในรูปของร้อยละ ทำให้ผู้เข้าร่วมโครงการเข้าใจได้ง่ายขึ้น

3. เปอร์เซ็นต์ไทล์ (Percentile) หรือ การจัดลำดับที่ (Rank)

การใช้ เปอร์เซ็นต์ไทล์หรือการจัดลำดับที่มีประโยชน์สำหรับข้อมูลที่กระจัดกระจายหรือมีลักษณะการกระจายแบบไม่ปกติ มีการตอบสนองที่แสดงลำดับ หรือเมื่อมีผลตอบสนองที่แตกต่างในจำนวนจำกัด แต่การใช้เทคนิคนี้ ควรใช้ด้วยความระมัดระวัง เนื่องจากการรายงานความสามารถของห้องปฏิบัติการ โดยการเรียงลำดับความสามารถไม่นิยมใช้ในการทดสอบความชำนาญ เพราะอาจจะทำให้เกิดความเข้าใจผิดแล้วนำไปสู่การแปลผลที่ผิด

4. Z-score

$$Z = \frac{x - X}{s}$$

โดยที่ s คือ การวัดหรือประมาณการกระจายที่เหมาะสม โดยเลือกให้ตรงกับข้อกำหนดของโครงการ สมการนี้สามารถใช้ได้ทั้งกรณีที่ X และ s เป็นผลการทดสอบที่ได้จากผู้เข้าร่วมโครงการ หรือ เมื่อ X และs ไม่ได้ใช้ผลการทดสอบจากผู้ร่วมโครงการ (ทั้งหมด) เช่น บางครั้งอาจจะใช้ค่าเบี่ยงเบนของ assigned value แทนก็ได้

ถ้าใช้ เทคนิค Z-score การประมาณ ค่าการกระจาย (s) ควรจะมีความน่าเชื่อถือ ทั้งนี้ขึ้นกับจำนวนข้อมูลที่มากเพียงพอเพื่อลดอิทธิพลจากค่าต่างสุด และทำให้ได้ค่าความไม่แน่นอนของการกระจายต่ำ

มีขั้นตอนทางสถิติหลายวิธีที่ใช้คำนวณค่า assigned values โดยใช้ผลการทดสอบของผู้เข้าร่วมโครงการ และค่าการกระจายของผลการทดสอบจากผู้เข้าร่วมโครงการ ซึ่งตามหลักสถิติประยุกต์สามารถแบ่งออกได้เป็น 2 กลุ่มใหญ่ๆ คือ Parametric statistics และ Non-parametric / Robust statistics[2]

4.1 สถิติแบบใช้พารามิเตอร์ (Parametric Statistics)

สถิติแบบใช้พารามิเตอร์ เป็นสถิติที่ใช้ข้อมูลเชิงปริมาณซึ่งรู้จักกันแพร่หลายและเป็นที่คุ้นเคยกันดีอยู่แล้ว โดยค่าเฉลี่ยเลขคณิตของข้อมูลจะถูกใช้ประมาณเป็นค่า assigned values และ ค่าเบี่ยงเบนมาตรฐานจะใช้แทนการวัดกระจายของข้อมูล สถิติแบบนี้จะใช้ได้กับผลการทดสอบที่มีการกระจายแบบปกติ (Normal distribution) แต่ถ้าข้อมูลที่ได้มีการกระจายแบบไม่ปกติ ค่าเฉลี่ยเลขคณิตและค่าเบี่ยงเบนมาตรฐานที่ใช้อธิบายผลการทดสอบจะไม่น่าเชื่อถือ ดังนั้นก่อนที่จะใช้สถิติแบบใช้พารามิเตอร์ จะต้องพิสูจน์ว่าข้อมูลมีการกระจายแบบปกติ ซึ่งสามารถทดสอบได้หลายวิธีเช่น Kolmogorov-Smirnov test, Lilliefors test เป็นต้น ลักษณะเด่นของสถิติแบบใช้พารามิเตอร์อย่างหนึ่งคือค่าต่างสุด (Extreme results) จะมีผลกระทบต่อค่าเฉลี่ยเลขคณิตและค่าเบี่ยงเบนมาตรฐาน ดังตัวอย่าง 1

ตัวอย่าง 1 : ผลของห้องปฏิบัติการที่เข้าร่วมโครงการจำนวน 20 ห้องปฏิบัติการ

ผลครั้งที่ 1 12.2 12.5 12.3 12.2 11.9 11.6 11.4 12.4 12.6 13.2 13.2 13.2 12.3 12.8 12.2 12.7 13.4 12.7 12.5 13.0

ค่าเฉลี่ย 12.52

ค่าเบี่ยงเบนมาตรฐาน 0.54

ผลครั้งที่ 2 122.0 12.5 12.3 12.2 11.9 11.6 11.4 12.4 12.6 13.2 13.2 13.2 12.3 12.8 12.2 12.7 13.4 12.7 12.5 13.0

ค่าเฉลี่ย 18.01

ค่าเบี่ยงเบนมาตรฐาน 24.48

จากตัวอย่าง พบว่าผลของห้องปฏิบัติการที่เข้าร่วมโครงการจำนวน 20 ห้องปฏิบัติการในครั้งที่ 1 จะได้ค่าเฉลี่ย 12.52 และค่าเบี่ยงเบนมาตรฐาน 0.54 แต่ถ้าผลของห้องปฏิบัติการหนึ่งรายงานผิดจาก 12.2 เป็น 122.0 (ผลครั้งที่ 2) จะได้ค่าเฉลี่ย 18.01 ซึ่งเพิ่มขึ้นถึง 44% และค่าเบี่ยงเบนมาตรฐาน 24.48 เพิ่มขึ้น 4000% จะเห็นได้ว่าค่าต่างสุด (extreme results) จะมีผลกระทบต่อค่าเฉลี่ยเลขคณิต

และค่าเบี่ยงเบนมาตรฐาน ดังนั้นการจะใช้สถิติแบบใช้พารามิเตอร์จำเป็นต้องทำการทดสอบ outlier ก่อน (โดยใช้การทดสอบของ Grubb's, Dixon's หรือ Cochrane's) เพื่อตัดอิทธิพลของข้อมูลที่เบี่ยงเบนก่อนที่จะคำนวณหาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐาน

4.2 สถิติแบบไม่ใช้พารามิเตอร์และสถิติแบบโรบัส (Non-parametric and robust statistics)

สถิติแบบไม่ใช้พารามิเตอร์และสถิติแบบโรบัสเป็นสถิติที่ใช้ข้อมูลที่อยู่ในรูปมาตราแสดงตำแหน่งหรือข้อมูลที่อยู่ในรูปลำดับที่ หรืออยู่ในรูปความถี่[1] และได้รับการยอมรับว่าสามารถใช้ได้กับข้อมูลที่มีการกระจายแบบไม่ปกติ ซึ่งสถิติแบบนี้มีประโยชน์กับโครงการทดสอบความชำนาญ เพราะจะพบบ่อยว่าข้อมูลหรือผลการทดสอบมีการกระจายแบบไม่ปกติ นอกจากนี้สถิติแบบไม่ใช้พารามิเตอร์ยังแตกต่างจากสถิติแบบใช้พารามิเตอร์เพราะข้อมูลที่มีค่าผิดปกติ (outliers) ไม่มีอิทธิพลต่อค่าเฉลี่ยของผลการทดสอบหรือการกระจายของผลทดสอบ ดังนั้นจึงไม่จำเป็นต้องใช้เทคนิคในการหา outlier ซึ่งเป็นข้อได้เปรียบของสถิติแบบนี้

ค่าเฉลี่ยที่ใช้ในสถิติแบบไม่ใช้พารามิเตอร์คือ มัธยฐานซึ่งเป็นจำนวนที่อยู่ตรงกลางเมื่อนำผลทดสอบมาเรียงจากน้อยไปมาก (ถ้าผลทดสอบมีจำนวนเป็นคู่ให้ใช้ค่าเฉลี่ยระหว่างค่าตรงกลางทั้งสอง) และค่าการกระจายของผลทดสอบในสถิติแบบไม่ใช้พารามิเตอร์คือ Median absolute deviation (MAD) เป็นค่ามัธยฐาน ของผลต่างสมบูรณ์ของผลทดสอบกับมัธยฐาน {MAD = median ($|x_i$- median|)} ถ้าจะทำให้มีค่าเท่ากับค่าเบี่ยงเบนมาตรฐานของการกระจายแบบปกติให้คูณ MAD ด้วย 1.483 ซึ่งค่านี้จะเรียกว่า MADe หรือ Inter-quatile range (IQR) เป็นผลต่างของค่าควอไทล์ที่ 3 กับค่าควอไทล์ที่ 1 (IQR = Q3-Q1) ถ้าจะทำให้มีค่าเท่ากับค่าเบี่ยงเบนมาตรฐานของการกระจายแบบปกติ ให้คูณ IQR ด้วย 0.7413 ซึ่งค่านี้จะเรียกว่า normalized IQR

ตัวอย่าง 2 : ผลของห้องปฏิบัติการที่เข้าร่วมโครงการจำนวน 20 ห้องปฏิบัติการ

ผลครั้งที่ 1 12.2 12.5 12.3 12.2 11.9 11.6 11.4 12.4 12.6 13.2 13.2 13.2 12.3 12.8 12.2 12.7 13.4 12.7 12.5 13.0

ค่ามัธยฐาน 12.50

MAD 0.30

IQR 0.75

ผลครั้งที่ 2 122.0 12.5 12.3 12.2 11.9 11.6 11.4 12.4 12.6 13.2 13.2 13.2 12.3 12.8 12.2 12.7 13.4 12.7 12.5 13.0

ค่ามัธยฐาน 12.55

MAD 0.35

IQR 0.925

จากตัวอย่าง พบว่าผลของห้องปฏิบัติการที่เข้าร่วมโครงการจำนวน 20 ห้องปฏิบัติการครั้งที่ 1 จะได้ ค่ามัธยฐาน 12.50, MAD 0.30, IQR 0.75 แต่ถ้าผลของห้องปฏิบัติการหนึ่งรายงานผิดจาก 12.2 เป็น 122.0 (ผลครั้งที่ 2) จะได้ค่ามัธยฐานเป็น 12.55 ซึ่งเพิ่มขึ้น 1% ค่า MAD 0.35 เพิ่มขึ้น 17 % และ IQR 0.925 เพิ่มขึ้น 23 % จะเห็นได้ว่าค่าต่างสุด (extreme results) จะมีผลกระทบต่อค่ามัธยฐาน ค่า MAD และค่า IQR น้อยกว่าค่าเฉลี่ยเลขคณิตและค่าเบี่ยงเบนมาตรฐานมาก

5. En numbers มักใช้กับโครงการเปรียบเทียบด้านการวัด (measurement comparison schemes) โดยมีสมการดังนี้

$$E_n = \frac{x - X}{\sqrt{U_{lab}^2 + U_{ref}^2}}$$

เมื่อ U_{lab} คือค่าความไม่แน่นอนของผลจากผู้เข้าร่วมโครงการ และ U_{ref} คือค่าความไม่แน่นอนของ assigned value จากห้องปฏิบัติการอ้างอิง

การประเมินความสามารถ (Evaluation of Performance)

ผลที่ได้จากการคำนวณความสามารถเชิงสถิติส่วนมากใช้ค่า Z-score และ En numbers ซึ่งมีเกณฑ์การตัดสินดังนี้

1. สำหรับ Z-score

ถ้าค่า $|Z| \leq 2$ แสดงว่าผลของห้องปฏิบัติการนั้นเป็นที่พอใจ

(satisfactory)

$2 < |Z| < 3$ แสดงว่าผลของห้องปฏิบัติการนั้นเป็นที่สงสัยต้องตรวจสอบ (questionable)

$|Z| \geq 3$ แสดงว่าผลของห้องปฏิบัติการนั้นไม่เป็นที่พอใจ (unsatisfactory)

2. สำหรับ En numbers

ถ้าค่า $|E_n| \leq 1$ แสดงว่าผลของห้องปฏิบัติการนั้นเป็นที่พอใจ (satisfactory)

$|E_n| > 1$ แสดงว่าผลของห้องปฏิบัติการนั้นไม่เป็นที่พอใจ (unsatisfactory)

ในการแปลผลจากผลการเข้าร่วมโครงการทดสอบความชำนาญ (Proficiency Testing, PT) มีขั้นตอนหลักๆ อยู่หลายขั้นตอนที่น่าสนใจ เพราะการทดสอบความชำนาญนั้นไม่ใช่พิจารณาแค่การ “ผ่าน” หรือ “ตก” จากการทดสอบในโครงการนั้นๆ แต่ยังมีอะไรอื่นที่น่าสนใจและน่าศึกษาจากผลที่ได้จากการเข้าร่วมโครงการนั้น ห้องปฏิบัติการที่มีผลการทดสอบทุกค่าให้ผลที่แสดงถึงความสามารถเป็นที่น่าพอใจไม่ได้หมายความว่าห้องปฏิบัติการนั้นเป็นห้องปฏิบัติการที่ดี แต่เป้าหมายที่ต้องการคือการคงความสามารถในระดับที่น่าพอใจนี้ไว้ หรืออีกนัยหนึ่งห้องปฏิบัติการที่มีผลการทดสอบไม่ดีก็ไม่ได้หมายความว่าเป็นห้องปฏิบัติการที่ต่ำกว่ามาตรฐาน แต่ผลที่ได้จากการเข้าร่วมโครงการควรจะนำมาศึกษาและแก้ไขปรับปรุงเพื่อที่จะไม่ให้เกิดซ้ำอีก

สรุปได้ว่าขั้นตอนหลักในทุกโครงการทดสอบความชำนาญ คือ การประเมินความสามารถของสมาชิกที่เข้าร่วมโครงการ โดยผู้ที่เป็นผู้จัดทำโครงการจะต้องมีการกำหนด 2 สิ่งคือ Assigned value ของตัวอย่างที่จะใช้ทดสอบ และช่วงการยอมรับ (Acceptable range) ของผลการทดสอบ สำหรับวิธีที่จะใช้ในการประเมินความสามารถของห้องปฏิบัติการนั้นมีหลายวิธี และไม่มีวิธีใดที่ใช้เป็นวิธีมาตรฐาน ขึ้นกับผู้จัดทำโครงการจะตัดสินใจและพัฒนาขั้นตอนสำหรับโครงการของตน อย่างไรก็ตามผู้จัดทำโครงการจะต้องบอกรายละเอียดถึงการเลือกใช้วิธีที่จะนำมาประเมินความสามารถของห้องปฏิบัติการแก่ผู้ที่จะเข้าร่วมโครงการทุกครั้ง

เอกสารอ้างอิง

Eurachem Nederland, task group :proficiency testing schemes and laboratory of the government chemist (LGC). Selection, use and interpretation of PT schemes. **Available: http:// www.eurachem.ul.pt.** 8 January 2003.

The International Organization for Standardization/The International Electrotechnical Commission. Proficiency testing by interlaboratory comparisons-Part 1: development and operation of proficiency testing schemes. **ISO/IEC Guide 43-1.** 1997.

กัลยา วานิชย์บัญชา. **การวิเคราะห์สถิติ: สถิติเพื่อการตัดสินใจ.** พิมพ์ครั้งที่ 5. กรุงเทพมหานคร: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย. 2544. หน้า 215-216.