

บทที่ 2

วรรณคดีที่เกี่ยวข้อง

การเสนอวรรณคดีที่เกี่ยวข้องในการวิจัยเรื่อง การวิเคราะห์เชิงประเมินค่าราคัศตรทางสถิติ ครั้งนี้ ผู้วิจัยได้ศึกษา เอกสาร และรายงานวิจัยที่เกี่ยวข้อง และนำเสนอสาระเป็น 5 ตอน ตอนแรก เป็นเรื่องความสำคัญของสถิติและค่าสถิติ ตอนที่ 2 เป็นเรื่องความหมายของการประเมิน หรือการปริทัศน์ค่าสถิติ ตอนที่ 3 เป็นเรื่องวิธีการประเมินค่าสถิติ ตอนที่ 4 การพัฒนาตัวบ่งชี้ (indicators) และ ตอนที่ 5 เป็นเรื่องแนวทางการประเมินค่าสถิติสำหรับการวิจัยครั้งนี้

ตอนที่ 1 ความสำคัญของสถิติและค่าสถิติ

ความสำคัญของสถิติทั้งในฐานะเป็นรายวิชา (discipline) เป็นวิทยาศาสตร์ (science) และเป็นวิชาชีพ (profession) นั้น เป็นที่ตระหนักดีในหมู่นักวิชาการ Bradley, R. A. (1982) กล่าวว่า ในฐานะรายวิชานั้น สถิติเป็นรายวิชาที่จำเป็นมากสำหรับการศึกษาระดับสูงโดยเฉพาะการศึกษาระดับบัณฑิตศึกษา Cockerill, R. และ Fried, B. (1991) ให้ความเห็นว่า นิสิตนักศึกษาที่เรียนในระดับอุดมศึกษาในสหรัฐอเมริกาทุกคน . ต้องผ่านการเรียนวิชาสถิติเป็นพื้นฐานอย่างน้อยวิชาหนึ่ง วิชาสถิติขั้นนำ (introductory statistics) จะเป็นรายวิชาที่มีผู้ลงทะเบียนเรียนเป็นจำนวนมากเสมอ เพราะความรู้ความเข้าใจที่ได้จากวิชาสถิติเป็นประโยชน์ต่อการอ่านรายงานวิจัย และการวิเคราะห์ข้อมูล

ความสำคัญของสถิติในฐานะเป็นวิทยาศาสตร์นั้น Bradley, R. A. (1982) สรุปว่า นักวิชาการมีความคิดเห็นแตกต่างกันไปหลายประเด็น แนวความคิดสำคัญมี 2 ประเด็น ประเด็นแรกนักวิชาการพิจารณาว่า สถิติเป็นวิทยาศาสตร์คณิตศาสตร์ (mathematical science) และสาขาวิชาสถิติในฐานะวิทยาศาสตร์คณิตศาสตร์นี้ สามารถอธิบายปรากฏการณ์ธรรมชาติได้ด้วยเนื้อหาสาระทฤษฎีของสถิติเอง เช่นเดียวกับวิทยาศาสตร์สาขาอื่นๆ ความคิดเห็นประเด็นที่ 2 นักวิชาการพิจารณาว่าสถิติเป็นวิทยาศาสตร์ประยุกต์ (applied science) ซึ่งไม่สามารถใช้ประโยชน์ได้โดยอิสระการใช้ประโยชน์สถิติต้องใช้ในฐานะเป็นเครื่องมือช่วยในการใช้ประโยชน์วิทยาศาสตร์สาขาอื่นๆ คุณประโยชน์ที่สำคัญ คือ การช่วยอธิบายปรากฏการณ์ที่ไม่แน่นอนตามหลักการ หรือ

กฎเกณฑ์ความไม่แน่นอน (law of uncertainty) อันเป็นที่มาของการใช้สถิติอนุมานในการอ้างอิงข้อเท็จจริงจากกลุ่มตัวอย่าง ไปสู่กลุ่มประชากร การใช้สถิติจึงจำต้องใช้ในลักษณะร่วมกับศาสตร์สาขาอื่นๆ

ความสำคัญของสถิติในฐานะวิชาชีพ เป็นที่ตระหนักกันดีในปัจจุบัน ตามมหาวิทยาลัยชั้นนำทั่วโลกมีการจัดการศึกษา เพื่อฝึกคนให้มีอาชีพเป็นนักสถิติ โดยมีโปรแกรมการเรียนการสอนถึงระดับปริญญาเอก Brogan, D. และ Kutner, M. H. (1986) สรุปว่า นักสถิติ (statistician) นอกจากจะประกอบอาชีพสอนวิชาสถิติแล้ว ยังประกอบอาชีพเป็นนักวิจัย ให้บริการใช้ระเบียบวิธีทางสถิติในการนำเสนอ จัดกระทำ วิเคราะห์ และแปลความหมายข้อมูล รวมทั้งเป็นนักวิชาการค้นคว้าวิชาการทางสถิติใหม่ๆด้วย Becker, B. J. (1996) กล่าวว่า วิธีการทางสถิติได้ถูกนำไปประยุกต์ใช้ในสาขาวิชาชีพต่างๆอย่างกว้างขวาง เมื่อใช้คำสำคัญ (key word) ว่า “statistics” ในการค้นรายการเอกสารโดยใช้ระบบการค้นเอกสารด้วยคอมพิวเตอร์ ของ ERIC พบรายการถึง 17,869 รายการ และเมื่อค้นบทความที่พิมพ์ในช่วง มกราคม 1993 ถึง ตุลาคม 1995 พบรายการบทความทางวิชาการที่เป็นบทความเรื่องสถิติรวม 15,623 รายการ เฉลี่ยเดือนละกว่า 700 รายการ Becker จึงสรุปว่าการรู้และการใช้สถิติ (statistical literacy) มีความสำคัญมากในปัจจุบัน

สำหรับสาขาทางการศึกษาหรือครุศาสตร์ สถิติมีความสำคัญทั้งในฐานะเป็นรายวิชาหนึ่งในหลักสูตร เป็นวิทยาศาสตร์ประยุกต์ และเป็นวิชาชีพ เช่นเดียวกับวิทยาศาสตร์สาขาอื่นๆ Cockerill, R. และ Fried, B. (1991) กล่าวว่า นักการศึกษาทุกคนที่ศึกษาในระดับอุดมศึกษาต้องเรียนรายวิชาทางสถิติอย่างน้อยหนึ่งวิชา เช่นเดียวกับนักวิชาการในสาขาอื่นๆ เนื่องจากนักการศึกษาใช้ระเบียบวิธีทางสถิติเป็นเครื่องมือในการแสวงหาความรู้โดยกระบวนการวิจัยทางการศึกษา วิชาสถิติที่เรียนในระดับอุดมศึกษา แบ่งออกได้เป็น 2 วิชา วิชาแรกคือ สถิติการศึกษา (educational statistics or statistics for education) อีกวิชาหนึ่ง มีลักษณะใกล้เคียงกัน คือ การศึกษาด้านสถิติ (statistical education) หรือการสอนสถิติ (teaching statistics) Harwell, M. R. และคณะ (1996) ให้คำจำกัดความว่า วิชาการศึกษาด้านสถิติ (statistical education) นอกจากจะครอบคลุมเทคนิควิธีทางสถิติที่เกี่ยวข้องกับการศึกษาแล้ว ยังครอบคลุมหัวข้อสำคัญเกี่ยวกับสถิติ 2 ประเด็น คือ หัวข้อสถิติอะไรบ้างที่ควรสอน และควรใช้วิธีการสอนแบบใดในการสอนหัวข้อดังกล่าว ส่วนวิชาสถิติการศึกษารอบคลุมเรื่องการใช้ระเบียบวิธีทางสถิติเป็นเครื่องมือแสวงหาความรู้ทางการศึกษา และครอบคลุมเรื่องการพัฒนาค่าสถิติและการศึกษาแนวโน้มการเปลี่ยนแปลงของระบบการศึกษา เมื่อเปรียบเทียบลักษณะวิชาสถิติการศึกษา กับ วิชาการศึกษาด้านสถิติ จะเห็นได้ว่าวิชาสถิติการศึกษาเป็นวิชาที่นักการศึกษาทุกคนต้องเรียน เพื่อใช้ประโยชน์ในฐานะเป็นเครื่องมือในการวิจัยแสวงหา

ความรู้ใหม่ แต่วิชาการศึกษาด้านสถิติเป็นวิชาที่นักการศึกษา ผู้ที่จะเป็นครูสอนวิชาสถิติเท่านั้นที่
ต้องเรียน

การเรียนการสอนวิชาสถิติไม่ว่าจะเป็นสถิติประยุกต์ สำหรับวิทยาศาสตร์สาขาใดๆ เช่นสถิติ
สำหรับนักจิตวิทยา สถิติสำหรับนักเศรษฐศาสตร์ สถิติสำหรับครูหรือนักการศึกษา ถ้าเป็นการสอน
สถิติประยุกต์เพื่อการวิจัยแล้ว ผู้เรียนต้องมีความรู้ทางระเบียบวิธีวิจัยมาก่อน หรือไม่เช่นนั้น ใน
วิชาสถิติประยุกต์นั้นต้องมีส่วนหนึ่งของวิชาว่าด้วยระเบียบวิธีวิจัย เป้าหมายของการสอนวิชาสถิติ
แตกต่างกันตามระดับและสาขาของผู้เรียน โดยทั่วไป เป้าหมายของการสอนวิชาสถิติอยู่ที่ การให้ผู้
เรียนมีความรู้ทางสถิติเพียงพอที่จะอ่าน และวิจารณ์งานวิจัยในสาขาของตนได้ สามารถทำงานวิจัย
ในส่วนหนึ่งของกิจกรรมการเรียนได้ และเมื่อสำเร็จการศึกษาแล้วสามารถทำการวิจัยในสาขางาน
อาชีพของตนได้ (Brogan, D. และ Kutner, M. H. , 1986)

สำหรับเนื้อหาสาระของวิชาสถิตินั้น Brogan, D. และ Kutner, M. H. (1986) ; Cockerill,
R. และ Fried, B. (1991) ; Belli, G. M. และ Seaver, W. L. (1989) สรุปว่า เนื้อหาสาระของ
วิชาสถิติแยกได้เป็น 2 วิชา คือ วิชาแรกได้แก่ วิชาสถิติเบื้องต้น มีเนื้อหาสาระครอบคลุมหัวข้อ
เทคนิคการวิเคราะห์ด้วยสถิติบรรยาย การแจกแจงความถี่ ทฤษฎีความน่าจะเป็น การแจกแจง
แบบไบนอมิยล และการแจกแจงปกติ การแจกแจงของค่าสถิติจากกลุ่มตัวอย่าง การประมาณค่า
พารามิเตอร์และการทดสอบสมมติฐาน ความคลาดเคลื่อนประเภทที่ I และ II ในการทดสอบ
สมมติฐาน การเปรียบเทียบประชากรกลุ่มเดียว สองกลุ่ม และมากกว่าสองกลุ่ม การวิเคราะห์
การถดถอยแบบง่าย และสถิตินันพารามตริก ส่วนวิชาที่สองหรือวิชาสถิติที่ผู้เรียนเรียนเป็นวิชา
ต่อจากสถิติเบื้องต้นนั้น มีเนื้อหาสาระครอบคลุมหัวข้อเทคนิควิธีทางสถิติขั้นสูง ได้แก่ โมเดล
เชิงเส้นทั่วไป (general linear model) เช่น การวิเคราะห์การถดถอยพหุคูณ การวิเคราะห์ความ
แปรปรวน และการวิเคราะห์ข้อมูลที่เป็นตัวแปรจัดประเภท และสถิติวิเคราะห์ตัวแปรพหุนาม
อย่างไรก็ดี เนื้อหาสาระของวิชาสถิติ ดังที่สรุปไว้นี้ มีความแตกต่างกันได้ตามลักษณะผู้เรียน
ผู้เรียนที่เรียนวิชาเอกทางสถิติอาจเรียนหัวข้อสำคัญบางหัวข้ออย่างลึกซึ้งซึ่งมากกว่าผู้เรียนที่เรียนวิชา
เอกอื่นที่ไม่ใช่สถิติ นอกจากนี้ ผู้เรียนที่ต้องการทำวิทยานิพนธ์ควรต้องเรียนรู้วิธีการใช้โปรแกรม
คอมพิวเตอร์สำเร็จรูปในการวิเคราะห์ข้อมูลทางสถิติด้วย ในขณะที่ผู้เรียนหลักสูตรที่ไม่ต้องทำ
วิทยานิพนธ์อาจเรียนรู้เพียงการอ่านผลการวิเคราะห์ข้อมูลจากคอมพิวเตอร์เท่านั้น

ด้านผู้สอนวิชาสถิติ Belli, G. M. และ Seaver, W. L. (1989) อธิบายว่า ผู้สอนวิชา
สถิติต้องใช้เทคนิควิธีหลายอย่างประกอบกัน เพราะสิ่งก้ำกัหรือความคิดรวบยอดทางสถิติ มี

ลักษณะเป็นนามธรรม ถ้าผู้เรียนไม่มีความเข้าใจ แต่เรียนรู้โดยการจำจะทำให้ผู้เรียนไม่สามารถใช้ประโยชน์จากสถิติได้ดีเท่าที่ควรจะเป็น ในการสอนผู้สอนต้องมีการยกตัวอย่าง การสาธิตด้วยภาพ อุปกรณ์ของจริง การทิสุงน์ ตลอดจนการประยุกต์ใช้ในสถานการณ์จริง ต้องมีการฝึกใช้โปรแกรมคอมพิวเตอร์ ให้ผู้เรียนทำแบบฝึกหัด และมีการอภิปรายด้วยอุปกรณ์การสอน สำหรับการสอนวิชาสถิติที่สำคัญที่จะช่วยให้ผู้เรียนเรียนรู้ได้รวดเร็วขึ้น คือ เอกสารประกอบการเรียน (handout) และตำรา

ตำราสถิติเป็นอุปกรณ์การสอนสำคัญชนิดหนึ่งในการสอนวิชาสถิติที่นักการศึกษาด้านสถิติให้ความสนใจมาก (Cobb, G. W. , 1987) Harwell, M. R. และคณะ (1996) กล่าวว่าปัจจุบันมีตำราสถิติอยู่เป็นจำนวนมาก แต่ผู้สอนวิชาสถิติมีปัญหาเกี่ยวกับการเลือกตำราที่มีความเหมาะสมกับจุดมุ่งหมายการสอน ซึ่งแตกต่างกันตามลักษณะผู้เรียน ลักษณะรายวิชา และให้ความสนใจกับการคัดเลือกตำราให้เหมาะสมกับการเรียนการสอน ความสนใจดังกล่าวเป็นที่มาของการคัดเลือกตำรา (textbook selection) ที่มีลักษณะตรงตามวัตถุประสงค์ของการทำงาน ซึ่งจะทำให้ได้โดยมีการประเมิน (evaluation) หรือการปริทัศน์ (review) ตำรา

ตอนที่ 2 ความหมายของการประเมินหรือการปริทัศน์ตำราสถิติ

ศาสตร์ไม่ว่าสาขาใดก็ตามที่มีการจัดการเรียนการสอนมาเป็นเวลานาน ผู้สอนย่อมแต่งตำรามาใช้ประกอบการเรียนการสอนให้ผู้เรียนได้ใช้ประโยชน์และเรียนรู้ได้เร็วขึ้น โดยไม่ต้องเสียเวลาจด ประเทศที่ยังมีการพัฒนาเท่าไร ตำราที่ใช้ในการเรียนการสอนก็ยังมีมากขึ้นเพียงนั้น ปัญหาที่ยุ่งยากมากในยุคปัจจุบัน คือ ปัญหาการตัดสินใจเลือกตำราที่เหมาะสมจากตำราที่มีอยู่หลากหลายทั้งในภาษาถิ่นและภาษาต่างประเทศ (Redei, G. P. , 1984) การคัดเลือกตำราเป็นงานหนัก และต้องอาศัยความรับผิดชอบอย่างสูงจากผู้มีอำนาจในการเลือกตำรา ผู้คัดเลือกต้องสร้างเกณฑ์ในการคัดเลือก ศึกษาตำราอย่างรอบคอบ แล้วนำมาพิจารณาตัดสินใจว่าตำรานั้นๆมีความเหมาะสมตรงตามเกณฑ์มากน้อยเพียงไร กระบวนการดำเนินการดังกล่าว คือการประเมินตำรา (Redei, G. P. , 1984 ; Hollabaugh, M. , 1989)

ในการจัดการเรียนการสอนสถิติในระดับอุดมศึกษา โดยเฉพาะระดับบัณฑิตศึกษา ผู้เรียนมีลักษณะหลากหลาย ในด้านภูมิหลัง ความสนใจพื้นความรู้ทางคณิตศาสตร์และงานวิจัยรวมทั้งมีวัตถุประสงค์ในการเรียนต่างกัน และผู้สอนมักพบว่าตำราเล่มหนึ่งดีในส่วนหนึ่งในขณะที่ตำราเล่มอื่นๆ มีดีในส่วนอื่นๆ (Harwell, M. R. และคณะ , 1996) จากสภาพปัญหาดังกล่าวจึงทำให้

นักการศึกษาผู้สอนวิชาสถิติ ให้ความสนใจศึกษาวิเคราะห์ประเมินคุณภาพของตำราหรือปริทัศน์ ตำราที่จะใช้ประกอบการเรียนการสอนกันมากขึ้น

การประเมิน (evaluation) ตำราที่จะใช้ประกอบการเรียนการสอน เป็นการวิเคราะห์เนื้อหาสาระของตำรา และพิจารณาตัดสินคุณค่าว่าตำรานั้นมีคุณภาพและปริมาณตรงตามความต้องการของผู้ใช้ตำราหรือไม่อย่างไร (Harwell, M. R. และคณะ , 1996) ทั้งนี้ผู้ประเมินต้องกำหนดประเด็นการประเมิน เช่น ความถูกต้องตามหลักวิชาของเนื้อหาสาระ ความสมบูรณ์และความลึกซึ้งของตำรา ความชัดเจนของการอธิบาย ความสมบูรณ์ของตัวอย่าง ความยากในการอ่าน เป็นต้น พวงปราง เพ็ญศรี (2528 อ้างถึงใน จุฬาลงกรณ์มหาวิทยาลัย , 2518) Harwell, M. R. และคณะ (1996) กล่าวว่า ในการประเมินคุณภาพตำรา ผู้ประเมินต้องมีเกณฑ์ในการประเมิน ผลการประเมินต้องให้สารสนเทศเกี่ยวกับตำราแก่ผู้บริโภคนในวงกว้าง ผลการประเมินต้องมีความเป็นธรรม และถูกต้องตามหลักวิชา เมื่อเปรียบเทียบการประเมินคุณภาพตำราสถิติ กับ การประเมินคุณภาพของแบบทดสอบหรือมาตรวัด หลักการประเมินควรเป็นแบบเดียวกัน

การปริทัศน์ (review) ตำรา มีความสำคัญเทียบเท่ากับการประเมินตำรา การปริทัศน์ตำรา หมายถึง การศึกษาวิเคราะห์เนื้อหาสาระของตำราในเชิงบรรยาย หรือเปรียบเทียบกับตำราอื่น (Huberty, C. J. และ Barton, R. M. , 1990 ; Farr, A. D. , 1981) หรือเสนอรายละเอียด ข้อมูลของตำราภายในเล่มว่าประกอบด้วยจำนวนหน้ากระดาษ จำนวนบท และความครอบคลุมเนื้อหาสำคัญในแต่ละบท (Simon, G. ,1996 ; Warde, D. W. , 1995 ; Ackerman, T. A. , 1993 ; Peck, R. ,1980) Farr, A. D. (1981) กล่าวว่า ผลการปริทัศน์ตำราต้องให้สารสนเทศเกี่ยวกับตำราแก่ผู้บริโภคเฉพาะกลุ่มเทคนิควิธี การปริทัศน์ตำรามีที่มาจากการปริทัศน์บทความหรืองานวิจัย โดยกลุ่มเพื่อน หรือที่เราเรียกว่า peer review ในวารสารทางวิชาการ การคัดเลือกบทความทางวิชาการ หรือรายงานการวิจัยลงพิมพ์ในวารสาร ใช้วิธีให้คณะบรรณาธิการทำหน้าที่ปริทัศน์ผลงานเหล่านั้น โดยให้คำวิจารณ์พร้อมทั้งข้อเสนอแนะเพื่อแก้ไข อย่างไรก็ตาม การปริทัศน์ตำรามีลักษณะแตกต่างออกไป เพราะเป็นการประเมินโดยบุคคลเพียงคนเดียวมากกว่าจะเป็นรูปคณะกรรมการ ปัจจุบันนี้จะพบรายงานการปริทัศน์ตำราในวารสารทางวิชาการ เพื่อแนะนำว่าตำรามีคุณภาพคุ้มค่ากับการซื้อ หรือการอ่านมากน้อยเพียงไร

ปัจจุบันนี้ในต่างประเทศโดยเฉพาะ สหรัฐอเมริกา มีผลงานวิชาการ ทั้งในรูปบทความ และ รายงานวิจัย ที่นำเสนอวิธีการและหลักเกณฑ์สำหรับการศึกษาวิเคราะห์เชิงประเมิน หรือการปริทัศน์ หรือการประเมินตำราสถิติเป็นจำนวนมาก Harwell, M. R. และคณะ (1996) สรุปในงาน

วิจัยประเมินตำราสถิติของเขาว่า มีผลงานทางวิชาการที่เป็นการประเมินตำราสถิติจำนวนมาก แต่การประเมินตำราเหล่านั้นขาดหลักทฤษฎีรองรับ ขาดเครื่องมือการประเมินที่มีคุณภาพ และมิได้ใช้วิธีการที่มีระบบ Harwell, M. R. และคณะ จึงได้เสนอแนวทางการวิเคราะห์เชิงประเมินตำราสถิติ โดยอาศัยผลงานวิชาการดังกล่าวเป็นพื้นฐาน และพัฒนาเครื่องมือสำหรับการประเมินมาใช้ประโยชน์ด้วย ผู้วิจัยจึงใช้ผลงาน ของ Harwell, M. R. และคณะ (1996) เป็นหลักประกอบกับผลงานวิจัยและเอกสารเกี่ยวกับการประเมินตำราสถิติหรือการปริทัศน์ตำราสถิติอื่นๆ ในการศึกษาเรื่องวิธีการวิเคราะห์เชิงประเมินตำราสถิติของนักวิจัยอื่นๆ ซึ่งจะได้นำเสนอในตอนต่อไป

ตอนที่ 3 วิธีการประเมินตำราสถิติ

จากการศึกษารายงาน เอกสาร และงานวิจัยที่เกี่ยวข้องกับการวิจัยครั้งนี้ ผู้วิจัยพบว่า เอกสารเกี่ยวกับการประเมินตำราสถิติเกือบทั้งหมด เป็นบทความวิชาการในวารสารจากต่างประเทศ และยังไม่พบในวารสารของประเทศไทย การประเมินตำราสถิติที่ผู้วิจัยศึกษา มีรูปแบบหลากหลาย ทั้งในด้านจุดมุ่งหมายการประเมิน และวิธีการประเมิน ดังที่ผู้วิจัยจะประมวลมาแนะนำแยกเป็น 4 ตอน คือ ตอนแรกเป็นเรื่องจุดมุ่งหมายของการประเมินตำราสถิติ ตอนที่ 2 เป็นเรื่องประเด็นการประเมินตำราสถิติ ตอนที่ 3 เป็นเรื่องวิธีดำเนินการประเมินตำราสถิติ และ ตอนที่ 4 เป็นเรื่องเกณฑ์การประเมินตำราสถิติ

3.1 จุดมุ่งหมายของการประเมินตำราสถิติ

จากการทบทวน เอกสาร และรายงานการวิจัยที่เกี่ยวข้อง ผู้วิจัยพบว่า มีผลงานวิชาการทั้งในรูปแบบบทความและรายงานการวิจัย เรื่อง การประเมินหรือการปริทัศน์ตำรา ซึ่งกำหนดจุดมุ่งหมายในการประเมินแตกต่างกัน สรุปแยกได้เป็น 3 ด้าน คือ

3.1.1 การประเมินตำราโดยมีจุดมุ่งหมายเพื่อคัดเลือกตำรามาใช้ประกอบการเรียนการสอน ผลงานวิชาการส่วนใหญ่ที่ผู้วิจัยศึกษาเป็นผลงานที่กำหนดจุดมุ่งหมายการประเมินเพื่อคัดเลือกตำรามาใช้ประกอบการเรียนรายวิชาต่างๆ ผู้วิจัยแยกประเภทผลงานวิชาการดังกล่าวเป็น 3 กลุ่ม กลุ่มที่หนึ่งเป็นผลงานวิชาการที่เป็นการประเมินตำราสถิติ กลุ่มที่สองเป็นผลงานวิชาการที่เป็นบทความเสนอสาระเกี่ยวกับหลักการในการประเมินตำราสถิติ และกลุ่มที่สามเป็นผลงานวิชาการที่เป็นการประเมินเพื่อคัดเลือกตำราวิชาอื่นๆ รายละเอียดแต่ละกลุ่มมีดังนี้

ผลงานวิชาการกลุ่มแรกที่เป็นการวิจัยประเมิน หรือ การวิเคราะห์ตำราสถิติ ได้แก่งานวิจัยของ Harwell, M. R. และคณะ (1996) ซึ่งประเมินตำราสถิติเพื่อสร้างแบบประเมินใช้ในการคัดเลือกตำรา งานวิจัยของ Simon, G. A. (1979) ซึ่งได้ประเมินตำราสถิติของ Freedman, D. ;

Pisani, R. และ Purves, R. , (1978) เพื่อแนะนำตำราให้นักวิชาการได้อ่านและเลือกใช้ งานวิจัยของ Cobb, G. W. (1987) ซึ่งได้ประเมินตำราสถิติ 16 เล่ม และจัดลำดับตำราสถิติตามความเหมาะสมที่จะใช้ในการเรียนการสอน และงานวิจัยของ Huberty, C. J. และ Barton, R. M. (1990) ซึ่งได้ประเมินตำราสถิติเรื่องการวิเคราะห์ตัวแปรพหุนาม โดยมีจุดมุ่งหมายเพื่อเลือกตำราไปใช้ประกอบการเรียนการสอน

ตัวอย่างการประเมินตำราสถิติ หรือการปริทัศน์ตำรา (book reviews) เพื่อคัดเลือกตำราที่มีความเหมาะสมกับการเรียนการสอนนั้น จะเห็นได้จากบทความภายใต้หัวข้อ การปริทัศน์ตำราตามวารสารวิชาการ ในรายงานวิจัยของ Huberty, C. J. และ Barton, R. M. (1990) ผู้วิจัยทั้งสองได้ให้ชื่อวารสารที่มีการปริทัศน์ตำราสถิติวิเคราะห์ตัวแปรพหุนาม เป็นแหล่งข้อมูลส่วนหนึ่งที่ผู้สนใจจะติดตามศึกษาวิธีการดำเนินการประเมินหรือปริทัศน์ตำราได้

กลุ่มที่สอง คือ ผลงานวิชาการที่เป็นบทความให้หลักเกณฑ์ในการประเมินตำรา ได้แก่บทความของ Redei, G. P. (1984) ซึ่งให้หลักเกณฑ์กว้างๆ ในการประเมินตำราทั่วไป เพื่อใช้ประกอบการคัดเลือกตำรา นอกจากนี้มีผลงานของ Belli, G. M. และ Seaver, W. L. (1989) ได้ให้ข้อเสนอแนะในการคัดเลือกตำราสถิติสำหรับนักเรียนนอกเวลา ตามแนวทางการวิจัยของ Cobb, G. W. (1987) เพื่อช่วยให้ผู้สอนได้คัดเลือกตำราที่มีคุณภาพมาใช้ประกอบการสอน

กลุ่มที่สาม คือ ผลงานวิชาการ ที่เสนอเป็นการประเมินเพื่อคัดเลือกตำราในวิชาอื่นๆนอกเหนือจากวิชาสถิติ ได้แก่ งานวิจัยของ Berry, K. และ Lee, D. (1982) ซึ่งเป็นการประเมินตำราวิทยาศาสตร์สาขาเคมี เพื่อคัดเลือกตำราไปประกอบการเรียนการสอนรายวิชา งานวิจัยของ Jacobson, M. D. (1965) ที่ได้ประเมินตำราวิชาฟิสิกส์และเคมี เพื่อคัดเลือกตำราเรียนระดับมัธยมศึกษาตอนปลายและอุดมศึกษา และงานวิจัยของ Hollabaugh, M. (1989) ซึ่งได้ประเมินตำราวิชาดาราศาสตร์ เพื่อคัดเลือกตำราใช้ในการเรียนของผู้เรียนระดับต่างๆ จนถึงระดับคุณวุฒิมัธยมศึกษา

3.1.2. การปริทัศน์ หรือประเมินตำราโดยมีจุดมุ่งหมายเพื่อเสนอแนะ นักวิชาการ นิสิต นักศึกษาเลือกซื้อหนังสือที่ดี มีคุณค่า และเหมาะสมกับราคา Farr, A. D. (1981) กล่าวไว้ในบทความเกี่ยวกับการปริทัศน์ตำรา (book reviews) ที่ลงพิมพ์ในวารสารทางวิชาการว่าผู้ปริทัศน์ต้องให้ข้อมูลเชิงประเมินว่า ตำรามีเนื้อหาและราคาเหมาะสมคุ้มค่ากับการใช้เงินซื้อหนังสือหรือคุ้มค่ากับการอ่านหรือไม่ อย่างไรก็ตามก็ตีหลักการพิจารณาที่ผู้ปริทัศน์ใช้ในการประเมิน เป็นแบบเดียวกับ การประเมินตำราหรือเอกสารที่ผู้ประเมินต้องใช้ในการตัดสินใจคัดเลือกตำราที่เหมาะสมมาใช้ประกอบการเรียนการสอน และเป็นหลักการคล้ายกับการคัดเลือกเอกสารที่เกี่ยวข้องกับการวิจัยมาทำรายงานเอกสารและการวิจัยที่เกี่ยวข้องนั่นเอง

3.1.3. การประเมินตำราโดยมีจุดมุ่งหมายเพื่อปรับปรุงตำราให้ดีขึ้น ในการจัดพิมพ์ครั้งต่อไป Farr, A. D. (1981) กล่าวไว้ในบทความเกี่ยวกับการปริทัศน์ตำราว่า การปริทัศน์หรือ

ประเมินตำรา นอกจากจะใช้ประโยชน์ในการคัดเลือกตำราประกอบการเรียนการสอนและใช้ประโยชน์ในการเลือกซื้อตำราแล้ว ยังใช้ประโยชน์ในการประเมินตำรา เพื่อปรับปรุงตำราให้ดีขึ้นด้วย ผู้วิจัยเห็นว่า หลักการพิจารณาการประเมินตำรา ที่ผู้ประเมินนำมาใช้ เป็นแบบเดียวกับการประเมินตำรา ที่ต้องใช้ในการตัดสินใจคัดเลือกตำราที่มีคุณภาพประกอบการเรียนการสอน ถ้าตำรานั้นมีผู้นิยมใช้สูง ความต้องการมาก การจัดพิมพ์ครั้งแรกอาจไม่เพียงพอต่อความต้องการ จำเป็นที่จะต้องมีการจัดพิมพ์ครั้งใหม่ขึ้น หากได้มีผู้ปริทัศน์หรือประเมินตำราเล่มนั้นไว้ ก็เป็นโอกาสที่ดีสำหรับผู้แต่งจะได้มีการแก้ไขปรับปรุงหนังสือเล่มนั้นให้ดีขึ้นต่อไป

จุดมุ่งหมายการประเมินตำราทั้ง 3 ด้านที่กล่าวมาแล้วนั้น แม้ว่าจะมีจุดมุ่งหมายต่างกัน แต่วิธีการประเมินมีลักษณะเช่นเดียวกัน และหลักการประเมินจะมีความคล้ายคลึงกัน ไม่ว่าจะทำเพื่อจุดมุ่งหมายข้อใด ผู้วิจัยพิจารณาเห็นว่าสำหรับนิสิต นักศึกษาผู้ใช้ตำราสถิติและผู้สอนวิชาสถิติ การประเมินตำราจะเป็นประโยชน์ในด้านการคัดเลือกตำราสถิติที่เหมาะสมใช้ประกอบการเรียนการสอน และประกอบการตัดสินใจเลือกซื้อตำราสถิติ ดังนั้นผู้วิจัยจึงสนใจประเมินตำราสถิติโดยกำหนดจุดมุ่งหมายเพื่อการคัดเลือกตำราประกอบการเรียนการสอนของผู้สอน และเพื่อการตัดสินใจของผู้สอนและนิสิตนักศึกษาในการเลือกซื้อหรือเลือกอ่านตำราสถิติ

3.2 ประเด็นการประเมินตำราสถิติ

จากการศึกษางานวิจัยเกี่ยวกับการประเมินตำราสถิติจำนวน 9 เรื่อง ผู้วิจัยพบว่าเป็นงานวิจัยซึ่งมีจุดมุ่งหมาย เพื่อคัดเลือกตำราสำหรับใช้ประกอบการเรียนการสอนวิชาสถิติมีจำนวน 6 เรื่อง ได้แก่ ผลงานวิจัยของ Simon, G. A. (1979) ; Belli, G. M. และ Seaver, W. L. (1989) ; Huberty, C. J. และ Barton, R. M. (1990) ; Redei, G. P. (1984) ; Cobb, G. W. (1987) และ Harwell, M. R. และคณะ (1996) ส่วนงานวิจัยอีก 3 เรื่องที่เหลือ เป็นงานวิจัยที่มีจุดมุ่งหมายเพื่อคัดเลือกตำราวิชาวิทยาศาสตร์สำหรับประกอบการเรียนการสอน โดยให้ความสำคัญกับการใช้ดัชนีความยากในการอ่านเป็นเกณฑ์ตัดสินใจในการคัดเลือกตำรา ได้แก่ ผลงานวิจัยของ Jacobson, M. D. (1965) ; Berry, K. และ Lee, D. (1982) และ Hollabaugh, M. (1989) ผลงานวิจัยทั้ง 9 เรื่องนี้ แต่ละเรื่องมีประเด็นการประเมินที่แตกต่างกัน และมีประเด็นการประเมินบางส่วนตรงกัน ในหัวข้อนี้ ผู้วิจัยนำเสนอประเด็นการประเมินจากการพิจารณาเครื่องมือที่ใช้ในการประเมินสำหรับงานวิจัยแต่ละเรื่อง ดังรายละเอียด ดังนี้

เรื่องที่ 1 Harwell, M. R. , Herrick, M. L. , Cutis, D. Mundfrom, D. และ Gold, K. (1966) ได้เสนอการประเมินคุณภาพตำราสถิติเบื้องต้น จำนวน 6 เล่ม เพื่อคัดเลือกตำราสำหรับครูผู้สอนและนิสิต นักศึกษา ใช้ประกอบการเรียนการสอนวิชาสถิติเบื้องต้น เครื่องมือการประเมินใช้แบบสอบถามชนิดเลือกตอบ และ มาตรฐานประเมินค่า กำหนดประเด็นในการประเมินตำราเป็น 8 ด้านคือ

ก. ระดับความยาก (Difficulty Level) การประเมินประเด็นนี้เป็นการวัดระดับความยากของตำรา โดยการใช้ข้อคำถามแบบมาตรฐานประเมินค่า 5 ระดับ จำนวน 3 ข้อ ได้แก่ ตำรามีระดับความยากในการอ่านมากน้อยอย่างไร อ่านตำราเข้าใจได้ง่ายมากน้อยระดับใด ตำรามีระดับความยากในการคำนวณมากน้อยอย่างไร

ข. ลักษณะของแบบฝึกหัด ในประเด็นนี้ เป็นการวัดลักษณะของแบบฝึกหัด โดยการใช้แบบสอบถามความคิดเห็นชนิดเลือกตอบ จำนวน 8 ตัวเลือก ได้แก่ โจทย์จากงานวิจัยจริง โจทย์สมมุติตามสถานการณ์จริง โจทย์ท้าทายความสามารถ โจทย์น่าสนใจ โจทย์ยากมาก โจทย์ง่ายเกินไป โจทย์ช่วยให้มีทักษะทางสถิติดีขึ้น ไม่มีโจทย์แบบฝึกหัด

ค. ดัชนีความน่าอ่าน (Readability Index) การประเมินประเด็นนี้ได้ใช้ดัชนีความน่าอ่านตามวิธีของ Hollabaugh, M. (1989) ซึ่งคำนวณจากค่าเฉลี่ยของดัชนีความน่าอ่านของตำราแต่ละหน้า จำนวน 6 หน้า ซึ่งมีค่าเท่ากับผลรวมของค่าเฉลี่ยจำนวนคำในแต่ละประโยค กับจำนวนคำที่มีค่ามากกว่า 3 พยางค์ในตำราแต่ละหน้า

ง. ความครอบคลุมของเนื้อหา (Content Coverage) การประเมินด้านการครอบคลุมของเนื้อหา นี้ ได้ใช้แบบสอบถามที่ใช้วัดความครอบคลุมของเนื้อหา เป็นข้อความให้ประเมินค่า 5 ระดับ ว่าผู้ตอบเห็นด้วยกับข้อความระดับใดจำนวน 3 ข้อ ได้แก่ ตำราครอบคลุมหัวข้อสถิติอย่างสมบูรณ์ หัวข้อมีความหลากหลายกว้างขวาง และ ตารางในแต่ละหัวข้อเป็นส่วนกับความสำคัญของหัวข้อนั้นมาก

จ. การนำเสนอเนื้อหา (Presentation) ในด้านการเสนอเนื้อหา นี้ ใช้แบบสอบถามที่ใช้วัดลักษณะการนำเสนอเนื้อหา เป็นข้อคำถามเลือกตอบ 2 ข้อ และข้อคำถามแบบประเมินค่า 5 ระดับ จำนวน 2 ข้อ ข้อคำถามแบบเลือกตอบข้อแรกถาม การเรียงลำดับหัวข้อในตำราสอดคล้องกับหัวข้อการสอนหรือไม่ ให้เลือกจาก 3 ตัวเลือก คือ สอดคล้องกัน มี 1 - 2 หัวข้อไม่สอดคล้องกัน และมี 3 หัวข้อขึ้นไปที่ไม่สอดคล้องกัน ข้อ 2 ถามการเชื่อมโยงระหว่างหัวข้อ ให้เลือก 3 ตัวเลือก คือ ทุกบทขึ้นอยู่กับทที่มาก่อนอย่างมาก ทุกบทขึ้นอยู่กับทที่มาก่อน และตารางในแต่ละบทสามารถนำเสนอได้โดยอิสระ ส่วนข้อคำถามแบบมาตรฐานประเมินค่าเป็นข้อคำถามให้ผู้ตอบตอบว่าเห็นด้วยกับข้อความระดับใด ได้แก่ ตำราแสดงวิธีการคำนวณที่จำเป็นสำหรับการใช้วิธีการทางสถิติ และ ตำราแสดงวิธีการคำนวณโดยเชื่อมโยงทฤษฎีและหลักการทางสถิติ

ฉ. การอธิบาย (Explanation) ในประเด็นการอธิบายนี้ ใช้แบบสอบถามที่ใช้วัดการอธิบายเป็นข้อคำถามแบบเลือกตอบจำนวน 6 ตัวเลือก ผู้ตอบเลือกคำตอบที่ตรงกับลักษณะการอธิบายเนื้อหาในตำราว่า ให้ความสำคัญในด้านใดต่อไปนี้คือ การยกตัวอย่าง การพิสูจน์ การบรรยาย การแสดงแผนภาพ การใช้กราฟ การอุปมาอุปไมย

ช. รูปแบบหรือลีลาการเขียน (Writing Style) การประเมินประเด็นนี้ใช้แบบสอบถามที่ใช้วัดรูปแบบหรือลีลาการเขียน เป็นข้อคำถามแบบเลือกตอบจำนวน 11 ตัวเลือก ผู้ตอบเลือกคำตอบที่ตรงกับรูปแบบหรือลีลาการเขียนตำราว่า เป็นแบบใดตามลักษณะต่อไปนี้ คือ ใช้ถ้อยคำรัดกุม ใช้คำพ้องเพื่อย ใช้ภาษาง่าย ยุ่งเหยิงไม่เป็นระบบ เหมาะสม เข้าใจง่าย ตรง รูปแบบไม่เป็นระบบ เป็นต้น

ซ. การประเมินในภาพรวม ในประเด็นการประเมินภาพรวมนี้ ได้ใช้แบบสอบถามความคิดเห็นผู้ประเมิน โดยใช้แบบสอบถามความเห็นชนิดเลือกตอบ 3 ข้อ จำนวน 3 ข้อ และแบบสอบถามแบบประเมินค่า 6 ระดับ จำนวน 3 ข้อ ได้แก่ จุดเน้นในการอธิบายตำรา รูปแบบการเขียนตำรา จุดเน้นภาพรวมทั้งหมด และความครอบคลุมของหัวข้อ

โดยสรุปประเด็นการประเมินของ Harwell, M. R. และคณะ (1996) เน้นความสำคัญของลักษณะเนื้อหาทั่วไปของตำราทั้งในด้านรูปแบบหรือลีลาการเขียน การอธิบาย การนำเสนอ ความน่าอ่าน และระดับความยากเป็นส่วนใหญ่ วิธีการประเมินครอบคลุมเรื่องแบบฝึกหัดและการประเมินในภาพรวมเป็นส่วนเพิ่มเติม โดยมีความสำคัญน้อยกว่าลักษณะเนื้อหาทั่วไป

เรื่องที่ 2. Redei, G. P. (1984) ได้รวบรวมประเด็นหรือเกณฑ์ที่ใช้ในการปริทัศน์ตำราต่างๆ เพื่อประเมินและคัดเลือกตำราประกอบการเรียนการสอน ประเด็นหรือเกณฑ์ที่ต้องการพิจารณาในการประเมินมี 5 ด้าน คือ

ก. ความเหมาะสมกับวิธีการสอนในรายวิชา (Course Tacting) มีประเด็นประเมิน 6 ประเด็น คือ เนื้อหาในตำราเป็นเรื่องที่ผู้สอนสามารถสอนได้อย่างดี เนื้อหาตรงตามความต้องการของผู้เรียนและไม่ซ้ำซ้อนกับวิชาอื่น พื้นความรู้ของผู้เรียนมีพอที่จะเข้าใจการอ่านเนื้อหาตำรา การจัดสัดส่วนเนื้อหาครอบคลุมสาระของวิชา การบูรณาการของเนื้อหาในตำราเหมาะสมและอธิบายทฤษฎีได้ชัดเจน และรูปแบบการเสนอเนื้อหาเหมาะสม

ข. ลักษณะเนื้อหาทั่วไป (General Characteristic) หรือความคิดเห็น (Point of View) เกี่ยวกับตำรา มีประเด็นประเมิน 8 ประเด็น คือ การเรียบเรียงเนื้อหาสอดคล้องกับจุดมุ่งหมาย การสอนจากง่ายไปยาก ตำรามีการอธิบายชัดเจน ตำรามีรายละเอียดลึกซึ้ง ตำรามีสัดส่วนที่สมดุลระหว่างการให้เหตุผลเชิงนิรนัยและอุปนัย ความยากง่ายในการอ่านเหมาะสมตำรามีรูปแบบ

การเสนอในเชิงวิทยาศาสตร์ การมีตัวอย่างและคำอธิบายชัดเจนและเพียงพอ และรูปลักษณ์ตำร่า น่าสนใจ

ก. ความถูกต้องของเนื้อหา (Accuracy of Content) มีประเด็นประเมิน 4 ประเด็น คือ ตำร่า มีปริมาณความคลาดเคลื่อนจากข้อเท็จจริงน้อย ตำร่ามีแหล่งอ้างอิงที่ถูกต้อง การมีคำอธิบาย ความเป็นมา การนำเสนอสาระมีความชัดเจนไม่สับสน

ง. ความทันสมัยของเนื้อหา (Currency of Content) มีประเด็นประเมิน 4 ประเด็น คือ ตำร่ามี สารสนเทศที่ทันสมัย ตำร่ามีสารสนเทศเก่าเพื่อเสนอในรูปประวัติความเป็นมาเท่านั้นตำร่ามีหัวข้อ ครอบคลุมสมบูรณ์ และตำร่ามีการเชื่อมโยงความรู้จากผลงานวิจัยดั้งเดิมกับความรู้ชายแดนจากการ วิจัย

จ. ภาพรวมของสารสนเทศ (Overall Information) มีประเด็นประเมิน 23 ประเด็น คือ ตำร่าระบุพื้นความรู้เดิม จำนวนหน้าของตำร่าเหมาะสมกับวิชา ตำร่ามีการประยุกต์ทฤษฎีอย่างมี ระบบ แบบฝึกหัดในตำร่ามีความเหมาะสม แบบฝึกหัดมีคำเฉลย มีคำอธิบายเฉลยคำตอบ แบบฝึกหัดมีความยากง่ายเหมาะสม แบบฝึกหัดมีการจัดเรียงลำดับเป็นขั้นตอน ตำร่ามีดัชนีที่เป็น ประโยชน์ในการค้นเรื่อง ตำร่ามีการอธิบายคำศัพท์ ตำร่ามีการสรุปท้ายบท มีแหล่งอ้างอิง เพียงพอ ตำร่ามีคำสำคัญ สัญลักษณ์และคำนิยามชัดเจน ตำร่ามีการอ้างอิงข้ามตอนให้ผู้อ่านติดตาม ศึกษาได้ ตำร่ามีการพิมพ์ที่ชัดและมีการพิมพ์เน้นโน้ตสำคัญ ตำร่ามีสารบัญเนื้อหา ตำร่า เสนอข้อมูลเชิงลึกหรือรายละเอียดในกรอบให้เห็นเด่นชัด ตำร่ามีคู่มือการสอนแก่ครู ในคู่มือครู มีแบบฝึกหัดเพิ่มเติมอย่างเหมาะสม ตำร่ามีคำถามสำหรับการสอบที่เหมาะสมกับการตรวจด้วย คอมพิวเตอร์ ตำร่ามีแหล่งอ้างอิงที่เหมาะสมกับการค้นคว้าต่อ ตำร่ามีภาคผนวกที่อำนวยความสะดวก ต่อการศึกษาต่อด้วยตนเอง และตำร่ามีราคาเหมาะสม

โดยสรุปประเด็นการประเมินของ Redei, G. P. ข้างต้นทั้ง 5 ประเด็น ครอบคลุมสาระทั้ง หมดยของตำร่า ทั้งในส่วนปลีกย่อยและภาพรวมของสารสนเทศ เพราะเป็นประเด็นการประเมินที่ ได้จากการบูรณาการประเมินของผู้ประเมินตำร่าหลายคน

เรื่องที่ 3. Cobb, G. W. (1987) ได้เสนอกรอบแนวทางการประเมินตำร่าสถิติเบื้องต้น จำนวน 16 เล่ม เพื่อใช้สำหรับประเมินคุณภาพของตำร่า ที่ใช้ประกอบการเรียนการสอนวิชา สถิติเบื้องต้น เครื่องมือประเมินใช้แบบสอบถาม กำหนดประเด็นในการประเมินตำร่าเป็น 3 ด้าน คือ

ก. ระดับวิชาการและคุณภาพของการอธิบายสาระ (Technical Level and Quality of Exposition) ประเด็นการประเมิน มี 2 ประเด็น คือ ประเด็นแรก คือ การใช้รูปแบบในการ

อธิบายมโนทัศน์ให้ชัดเจนและการเชื่อมโยงมโนทัศน์ ประเด็นที่ สอง คือ การเน้นหลักพิชคณิต ในการอธิบายความเกี่ยวข้องระหว่างมโนทัศน์ ประเด็นทั้งสองเป็นการประเมินครอบคลุมหัวข้อ มาตรฐาน 4 เรื่องคือ ค่าเฉลี่ยของกลุ่มตัวอย่าง (sample mean) ค่าเบี่ยงเบนมาตรฐาน (sample standard deviation , S.D.) การแจกแจงปกติ (normal distribution) การแจกแจงของค่าเฉลี่ย จากกลุ่มตัวอย่าง (sampling distribution of mean)

ข. ความครอบคลุมของหัวข้อ (Topic Coverage) การประเมินในด้านนี้เป็นการเปรียบเทียบ ตำรา กับเนื้อหาวิชาสถิติเบื้องต้น ได้แก่ การถดถอย (regression) การวิเคราะห์ความแปรปรวน (analysis of variance) การวิเคราะห์ข้อมูลเพื่อสำรวจ (exploratory data analysis) ผลการวิเคราะห์ ด้วยคอมพิวเตอร์ (computers output) ความน่าจะเป็น (probability) สถิตินั้นพารามเมตริก (nonparametric statistics) และลักษณะสำคัญอื่นๆ เช่น การมีฐานข้อมูลให้เป็นตัวอย่าง การมี โครงการวิจัยตัวอย่าง เป็นต้น

ค. คุณภาพของแบบฝึกหัด (Qualities of Exercises) การประเมินในด้านนี้เป็นการพิจารณา ใน 3 ประเด็น ได้แก่ แบบฝึกหัดใช้ข้อมูลที่เป็นจริง การหาคำตอบแบบฝึกหัดนำไปสู่คำถามที่น่าสนใจ แบบฝึกหัดต้องมีสัดส่วนการใช้ความคิดมากกว่าความจำ ในแต่ละด้าน Cobb, G. W. ยังพิจารณาประเมินว่า ตำรามีเนื้อหาครบตามหัวข้อย่อยในแต่ละเรื่องหรือไม่ด้วย เช่น ในหัวข้อ การวิเคราะห์ความแปรปรวน ต้องมีหัวข้อย่อยในเรื่องต่อไป คือ การวิเคราะห์ความแปรปรวนทาง เดียว การวิเคราะห์ความแปรปรวน 2 ทาง เมื่อขนาดของกลุ่มตัวอย่างแต่ละเซลล์เป็น 1 และ มากกว่า 1 การวิเคราะห์เปรียบเทียบรายคู่ การวิเคราะห์ปฏิสัมพันธ์ เป็นต้น

ประเด็นการประเมินของ Cobb, G. W. เน้นความสำคัญที่เนื้อหาสาระทางสถิติเป็นส่วน ใหญ่ วิธีการประเมินเน้นความถูกต้องและความชัดเจน ตลอดจนความครอบคลุมในการเสนอ เนื้อหาของสถิติ

เรื่องที่ 4. Huberty, C. J. และ Barton, R. M. (1990) ได้ประเมินคุณภาพตำรา การวิเคราะห์ ตัวแปรพหุนาม จำนวน 15 เล่ม เพื่อคัดเลือกตำราสำหรับใช้ประกอบการเรียนการสอนวิชาสถิติ เครื่องมือการประเมินใช้มาตรประเมินค่าแบบ Likert กำหนดประเด็นในการประเมินตำราเป็น 4 ด้าน คือ ความครอบคลุมของเนื้อหา การแนะนำให้ลงมือทำ ความน่าอ่าน และคุณภาพของ แบบฝึกหัด ดังรายละเอียดต่อไปนี้

ก. ความครอบคลุมของเนื้อหา (Content Coverage) การประเมินประเด็นนี้เป็นการประเมิน เฉพาะคุณภาพของตำรา โดยแยกพิจารณา 3 ด้าน คือ ความถูกต้อง (correctness) ความทันสมัย (currentness) และ ความสมบูรณ์ (thoroughness) ในแต่ละด้านเป็นการประเมินใน 4 เรื่อง คือ

เรื่องคำอธิบายมโนทัศน์ทางสถิติและการคำนวณ เรื่องการใช้สัญลักษณ์และศัพท์ เรื่องการให้คำนิยาม และเรื่องการวิเคราะห์และอภิปรายผลการวิเคราะห์ การรวบรวมข้อมูลใช้แบบสอบถามแบบมาตราประเมินค่า 5 ระดับ ซึ่งจะมีข้อคำถาม 3 ด้าน ด้านละ 4 เรื่อง รวม 12 ข้อ

ข. การแนะนำให้ลงมือทำ (How to do) ในประเด็นนี้เป็นการวัดการเรียนรู้จากการลงมือทำตามคำอธิบายในตำราในเรื่องการวิเคราะห์ตัวแปรพหุคูณและการอภิปรายผล โดยการใช้คำถามแบบมาตราประเมินค่า 5 ระดับ จำนวน 3 ประเด็น คือ การแนะนำการใช้โปรแกรมสำเร็จรูปทางสถิติที่ทันสมัย การแนะนำการเลือกใช้โปรแกรมคอมพิวเตอร์ที่เหมาะสม และการอภิปรายการเลือกใช้สถิติ

ค. ความน่าอ่าน (Readability) การประเมินด้านความน่าอ่านเป็นการเปรียบเทียบเนื้อหาในตำราว่าผู้อ่านที่เป็นผู้เรียนอ่านเข้าใจหรือไม่ โดยเทียบกับตำราสถิติของ Pedhazur ว่าเป็นตำราที่อ่านเข้าใจง่าย การรวบรวมข้อมูลทำแบบสอบถามแบบมาตราประเมินค่า 5 ระดับ จำนวน 2 ข้อ ได้แก่ เนื้อหาของตำราอ่านเข้าใจง่าย และการนำเสนอสูตรใช้หลักพีชคณิตเมทริกซ์

ง. คุณภาพของแบบฝึกหัด (Quality of Exercises) เนื่องจากแบบฝึกหัดจะช่วยให้ผู้เรียนหัดการเรียนรู้และมีทักษะที่ถูกต้อง การประเมินคุณภาพแบบฝึกหัดทำได้โดยการใช้แบบสอบถามแบบมาตราประเมินค่า 5 ระดับ จำนวน 5 ข้อ ได้แก่ แบบฝึกหัดใช้ฐานข้อมูลจากการวิจัยจริง มีการเฉลยคำตอบของแบบฝึกหัด แบบฝึกหัดในตำราครอบคลุมการแปลความหมาย การอภิปราย และมีโจทย์ปัญหาให้ผู้เรียนใช้ความคิด และโจทย์ให้คำนวณ

โดยสรุปประเด็นการประเมินของ Huberty, C. J. และ Barton, R. M. ได้เน้นความสำคัญคุณภาพของตำราในด้าน ความครอบคลุมของเนื้อหา การแนะนำให้ลงมือทำ ความน่าอ่าน และแบบฝึกหัด ซึ่งคล้ายคลึงกับประเด็นการประเมินของ Cobb, G. W.

เรื่องที่ 5. Belli, G. M. และ Seaver, W. L. (1989) ได้ประเมินตำราสถิติสำหรับนักศึกษาภาคนอกเวลาที่เรียนด้วยตนเองนอกมหาวิทยาลัยโดยใช้วิธีการตามแบบของ Cobb, G. W. (1987) มาปรับปรุงให้มีประเด็นคำถามเกี่ยวกับ ความลึกซึ้ง แบบฝึกหัด และ การใช้คอมพิวเตอร์ ผลการประเมินสรุปว่า เกณฑ์ในการเลือกตำราสำหรับนักศึกษากลุ่มนี้ แม้จะมีหลักการคล้ายกับแบบเลือกตำราสถิติสำหรับนักศึกษาภาคปกติ แต่จะมีข้อแตกต่างกันเล็กน้อย เพราะนักศึกษากลุ่มนี้มีภูมิหลังทางคณิตศาสตร์ต่างกัน มีเวลาให้การเรียนรู้เพราะต้องทำงาน แต่มีแรงจูงใจในการเรียนสูง ลักษณะที่แตกต่างกัน ได้แก่ จุดมุ่งหมายของการใช้ตำรา เพราะนักศึกษากลุ่มนี้มีเวลาเข้าห้องสมุดน้อย จึงต้องเรียนโดยอิงตำราเป็นส่วนใหญ่ การคัดเลือกตำราจึงต้องเลือกตำราที่เนื้อหามีความ

ถึกซึ่งและครอบคลุมหัวข้อกว้างขวาง มีแบบฝึกหัดที่เชื่อมโยงกับการบรรยายของผู้สอน ในส่วนของการวิเคราะห์ข้อมูลด้วยคอมพิวเตอร์ต้องมีผลการวิเคราะห์เสนอไว้ในตำราด้วย

โดยสรุปวิธีของ Belli, G. M. และ Scaver, W. L. เสนอให้ใช้วิธีของ Cobb, G. W. แต่ให้เพิ่มรายละเอียดอีก 3 ข้อ คือ ตำรามีเนื้อหาถึกซึ่งและครอบคลุมหัวข้อกว้างขวาง มีแบบฝึกหัดที่เชื่อมโยงกับการบรรยายของผู้สอน และมีการเสนอผลการวิเคราะห์ข้อมูลด้วยคอมพิวเตอร์ในตำรา

เรื่องที่ 6. Simon, G. A. (1979) ได้ปริทัศน์ตำราสถิติของ Freedman, D. ; Pisani, R. และ Purves, R. (1978) เพื่อเสนอแนะให้นักวิชาการได้อ่านและเลือกใช้ โดยมีประเด็นการปริทัศน์ดังนี้ การเรียงลำดับของหัวข้อ วิธีการนำเสนอที่เข้าใจง่าย การบอกขั้นตอนการหาค่าสถิติโดยไม่ใช้สูตรแต่ใช้ภาษาเขียน การมีตัวอย่างที่เข้าใจง่าย การมีส่วนที่เป็นนวัตกรรมใหม่ๆ การมีแบบฝึกหัดให้นิสิตกำนวณเองพร้อมคำตอบ การมีเชิงอรรถอธิบายการหาค่าสถิติด้วยตัววิธีอื่นๆ การมีแบบสอบวัดความรู้พื้นฐานก่อนเรียน การมีคู่มือครู นอกจากนี้ ผู้วิจารณ์ได้สอบถามนิสิตกที่ใช้ตำราเล่มนี้ในการเรียน ว่า อินดิชยตำราเล่มนี้ในราคา 5 เหรียญอเมริกันหรือไม่ ปรากฏว่านิสิตกิน 50 % ต้องการเก็บตำราไว้ใช้ต่อไป

เรื่องที่ 7. Berry, K. และ Lee, D. (1982) ได้เสนอประเด็นการประเมินคุณภาพตำราวิทยาศาสตร์ เพื่อคัดเลือกตำราสำหรับผู้สอนและและผู้เรียนใช้ประกอบการเรียนการสอน วิชาเคมีรวม 9 ประเด็น และให้ข้อเสนอแนะว่า การประเมินคุณภาพตำราทั้ง 9 ประเด็นนี้ อาจใช้เครื่องมือวัดที่เป็นแบบสอบถามรวบรวมข้อมูล ในการประเมินตำราต่างๆ ไปได้ ประเด็นการประเมิน 9 ประเด็น มีดังต่อไปนี้

ก. ความน่าอ่าน (Readability) ประเด็นการประเมินนี้คือการวัดระดับความยากในการอ่านตำรา โดยวัดเป็นดัชนีความน่าอ่าน ในรูปการจัดอันดับหรือคุณค่า เช่น การใช้ดัชนีความน่าอ่านของ Fry ซึ่งวัดเป็นอันดับและระดับการอ่านที่แท้จริงของผู้เรียน

ข. ความเป็นผู้ทรงความรู้ (Authority) ของผู้แต่ง ประเด็นการประเมินนี้ แยกเป็นประเด็นด้านผู้แต่ง และด้านตำรา ได้แก่ ผู้แต่งมีคุณวุฒิและมีความเชื่อถือได้ และตำรามีสาระที่ทันสมัย

ค. คำศัพท์ (Vocabulary) ประเด็นการประเมินคุณภาพตำราในด้านคำศัพท์ แยกเป็น 3 ประเด็นย่อย คือ การให้ความหมายคำศัพท์สำคัญในตำรา การพิมพ์คำศัพท์สำคัญด้วยตัวหนาหรือตัวเอนเพื่อให้จำได้ง่าย และการมีคำนิยามที่ชัดเจนอ่านเข้าใจง่าย

ง. มโนทัศน์ (Concept) ประเด็นการประเมินด้านมโนทัศน์ของเนื้อหาวิชา มี 7 ข้อ ได้แก่ การเสนอมโนทัศน์ตรงตามที่อยู่เรียนคาดหวัง มโนทัศน์มีความถูกต้องและเป็นปรนัย การเสนอ มโนทัศน์มีลำดับเหมาะสมกับแผนการสอน รูปแบบของตำราแยกมโนทัศน์สำคัญชัดเจน มีคำอธิบายมโนทัศน์ให้ผู้อ่านเข้าใจ มีคู่มือการสอนหรือมีแบบฝึกหัด การเสนอมโนทัศน์กระชับและ ไม่สับสน

จ. การเสนอเนื้อหาของตำรา (Presentation of material) ประเด็นการประเมินด้านการนำ เสนอเนื้อหา มี 2 ข้อ ได้แก่ มีบรรณานุกรมที่ต้องการให้ค้นคว้าเพิ่มเติมท้ายบทหรือท้ายเล่ม และเนื้อหาเกี่ยวข้องกับชีวิตประจำวันและวิชาอื่นๆ

ฉ. แบบฝึกหัดหรือคู่มือแนะนำการเรียน (Exercise or study Guide) ประเด็นการประเมิน ด้านแบบฝึกหัดมี 3 ข้อ ได้แก่ แบบฝึกหัดมีคุณค่า แบบฝึกหัดช่วยสร้างเสริมการเรียนรู้คำศัพท์ และมโนทัศน์ และโจทย์แบบฝึกหัดกระตุ้นให้เกิดการคิด

ช. ทัศนูปกรณ์ (Visual Aids) ประเด็นการประเมินด้านทัศนูปกรณ์ช่วยให้อ่านตำราเกิด ความเข้าใจดีขึ้น มีจำนวน 3 ข้อ ได้แก่ การใช้กราฟเส้น แผนภูมิ อื่นๆเสริมความเข้าใจ มี ตัวอย่างชัดเจนและมีความหมาย มีตัวอย่างชัดเจนและถูกต้อง

ซ. ส่วนประกอบของตำรา (Part of Textbook) ประเด็นการประเมินด้านนี้ แยกเป็น 4 ข้อ ได้แก่ การมีสารบัญที่เป็นประโยชน์ การมีดัชนีที่สมบูรณ์ การมีเฉลยแบบฝึกหัด และมีคู่มือ แนะนำการเรียน

ณ. คุณภาพและราคาของตำรา (Quality and Cost) ประเด็นการประเมินด้านนี้ มีจำนวน 5 ข้อ ได้แก่ การพิมพ์ชัดเจนและน่าอ่าน รูปเล่มตำราแข็งแรง คุณภาพและสีกระดาษเหมาะสม มีราคาเหมาะสม และเป็นตำราคุ้มค่าสมราคา

โดยสรุปประเด็นการประเมินของ Berry, K. และ Lee, D. ได้เสนอประเด็นการประเมิน นอกเหนือจากที่ผู้ประเมินคนอื่นทำไว้เกี่ยวกับสาระของตำรา คือ ความเป็นผู้ทรงความรู้ คำศัพท์ มโนทัศน์ การเสนอเนื้อหา ทัศนูปกรณ์ ส่วนประกอบของตำรา และคุณภาพและราคา

เรื่องที่ 8. Jacobson, M. D. (1965) ได้เสนอการประเมินตำราฟิสิกส์และเคมี จำนวน 13 เล่ม เพื่อคัดเลือกตำราสำหรับประกอบการเรียนการสอนวิชาฟิสิกส์และเคมี โดยใช้ดัชนีความยากใน การอ่าน (reading difficulty index) เป็นประเด็นการประเมิน เครื่องมือประเมินเป็นแบบสอบขีด เส้นใต้ (underlining test) ในการรวบรวมข้อมูล ผู้เรียนจะต้องอ่านข้อความที่สุ่มมาจากตำราและ ขีดเส้นใต้คำศัพท์ที่ยากอ่านไม่เข้าใจ ดัชนีความยากในการอ่านคือค่าเฉลี่ยจำนวนคำที่ขีดเส้นใต้ Jacobson, M. D. ได้หาค่าความเที่ยงของดัชนีความยากในการอ่านแบบสอบซ้ำ ได้ค่า ความเที่ยง

เท่ากับ 0.95 และ 0.85 สำหรับวิชาฟิสิกส์และเคมีตามลำดับ นอกจากนี้เขายังได้หาค่าความตรงร่วมสมัยได้ค่าความสัมพันธ์ทางลบ ระหว่างดัชนีความยากในการอ่านกับจำนวนคำศัพท์ที่ผู้เรียนมีความรู้ ดัชนีความยากในการอ่านมีความสัมพันธ์กับตัวทำนาย รวม 5 ตัวแปร ได้แก่ จำนวนคำต่ออนุประโยค จำนวนคำศัพท์คณิตศาสตร์ที่ยาก คำศัพท์ที่ยากสูงกว่าระดับ 6,000 คำ ในรายการศัพท์ของ Thomdike 20,000 จำนวนคำศัพท์วิทยาศาสตร์ จำนวนคำที่มี 3 พยางค์ขึ้นไปต่อจำนวนคำในข้อความที่ให้อ่าน

เรื่องที่ 9. Hollabaugh, M. (1989) ได้ประเมินตำราดาราศาสตร์ จำนวน 5 เล่มเพื่อคัดเลือกตำราสำหรับผู้สอนและผู้เรียนใช้ประกอบการเรียนการสอนวิชาดาราศาสตร์ ประเด็นการประเมิน คือ ดัชนีความยากในการอ่าน (reading difficulty) ของ Gunning ที่เรียกว่า ดัชนีฟ็อก (fog index) หรือ ดัชนีหมอก หรือดัชนีความไม่ชัดเจน ซึ่งคำนวณจาก จำนวนประโยค จำนวนคำต่อประโยค จำนวนคำที่มี 3 พยางค์ขึ้นไปในข้อความที่มี 100 คำ ถ้าดัชนีหมอกหรือดัชนีความไม่ชัดเจนมีค่าสูง แสดงว่าตำราอ่านยาก สูตรการคำนวณของดัชนีหมอกหรือดัชนีความไม่ชัดเจน มีดังนี้

$$\text{ดัชนีหมอกหรือดัชนีความไม่ชัดเจน} = 0.4 \left(\text{ค่าเฉลี่ยจำนวนคำต่อประโยค} + \text{จำนวนคำ 3 พยางค์ขึ้นไป} \right)$$

นอกจากนี้ Hollabaugh, M. ยังได้กล่าวถึงองค์ประกอบอื่นๆที่มีผลต่อการคัดเลือกตำรา ซึ่งได้แก่ ความลึกของเนื้อหา ความครอบคลุมหัวข้อ ระดับการคำนวณ การอธิบายที่เหมาะสม รูปแบบการเขียน และการใช้ศัพท์ทางวิทยาศาสตร์

เมื่อเปรียบเทียบประเด็นการประเมินตำราที่กล่าวมาข้างต้น พบว่าประเด็นการประเมินมีความคล้ายคลึงกัน ดังตารางที่ 1 ผู้วิจัยจัดกลุ่มประเด็นการประเมินที่มีความคล้ายคลึงกันอยู่รวมกันได้เป็น 11 ด้าน ประเด็นการประเมินทั้ง 11 ด้าน ได้แก่ ลักษณะรูปเล่ม รูปแบบการเขียน ความน่าอ่านของตำรา วิธีการนำเสนอเนื้อหา ความครอบคลุมของเนื้อหา ความถูกต้องของเนื้อหา ตารางที่ทำให้เข้าใจดีขึ้น ลักษณะของแบบฝึกหัด ส่วนเสริมของตำรา การให้คำแนะนำในการสอน และลักษณะของผู้แต่ง ดังนี้คือ

1. ลักษณะรูปเล่ม การประเมินด้านลักษณะรูปเล่มนี้ มีประเด็นการประเมิน 9 ประเด็น คือ รูปลักษณะของตำร่าน่าสนใจ (Redei, G. P. , 1984) รูปเล่มของตำราแข็งแรง (Berry, K. และ Lee, D. , 1982) การมีจำนวนหน้าเหมาะสมกับวิชา (Redei, G. P. , 1984) มีการพิมพ์ที่ชัดเจน น่าอ่าน (Redei, G. P. , 1984 ; Berry, K และ Lee, D. , 1982) มีการพิมพ์ที่เน้นมโนทัศน์ที่สำคัญ (Redei, G. P. , 1984) มีการพิมพ์คำศัพท์ที่สำคัญด้วยตัวเอนหรือตัวหนา (Berry, K และ Lee, D. , 1982) การมีสีกระดาษเหมาะสม (Berry, K และ Lee, D. , 1982) การมีราคา

ตารางที่ 1 สรุปประเด็นการประเมินจากงานวิจัยประเมินตำราสถิติ

ประเด็นการประเมิน	Har	Re	Cobb	H&B	B&S	Slim	B&L	Jac	H
1. ลักษณะรูปเล่ม									
รูปเล่มของตำร่าน่าสนใจ		*							
รูปเล่มของตำร่าแข็งแรง							*		
มีจำนวนหน้าเหมาะสมกับวิชา		*							
มีการพิมพ์ที่ชัดเจนน่าอ่าน		*					*		
มีการพิมพ์เน้นมโนทัศน์ที่สำคัญ		*							
มีการพิมพ์คำศัพท์ที่สำคัญด้วยตัวหนาหรือตัวหนา							*		
มีสีกระดาษเหมาะสม							*		
มีราคาที่เหมาะสม		*					*		
ความมีคุณภาพของตำร่าที่ดี							*		
2. รูปแบบการเขียน									
วิธีการเขียนถูกต้องตามหลักภาษา	*								
การใช้ถ้อยคำรัดกุมไม่ฟุ่มเฟือย	*								
การใช้ภาษาที่ง่าย	*								
การเขียนไม่เป็นระบบ ไม่มีระบบในการเขียน	*								
มีการเขียนที่เหมาะสมกับระดับความรู้ของผู้เรียน	*								
วิธีการเขียนอ่านแล้วเข้าใจง่าย	*				*				
มีการเขียนที่ไม่ซ้ำประเด็นสำคัญของเรื่อง	*								
มีการเขียนที่ใช้ภาษาสะอาดสวย	*								
มีการเขียนเชิงอุปมา-อุปนัย	*								
มีการเขียนเป็นแบบบรรยาย	*								
3. ความน่าอ่านของตำร่า									
ระดับความยากของตำร่า	*			*			*		
ระดับความยากในการอ่านตำร่า	*	*							
ระดับความถุกซึ่งทางวิชาการ			*						
ระดับความยากในการคำนวณ	*								*
ระดับการใช้คำศัพท์ที่ยาก									*
ดัชนีความยากในการอ่าน / ดัชนีความน่าอ่าน	*			*				*	*

ตารางที่ 1 (ต่อ)

ประเด็นการประเมิน	Har	Re	Cobb	H&B	B&S	Slim	B&L	Jac	Hol
4. วิธีการนำเสนอเนื้อหา									
วิธีการนำเสนอเนื้อหาเรียงจากง่ายไปยาก	*								
การเรียงลำดับหัวข้อสอดคล้องกับหัวข้อการสอน	*					*			
การเสนอสาระมีความชัดเจน		*							
มีคำอธิบายบนโน้ตบุ๊คให้ผู้อ่านเข้าใจง่าย							*		
การเสนอบนโน้ตบุ๊คมีความกระชับ ไม่สับสน							*		
มีการเชื่อมโยงเนื้อหาระหว่างหัวข้อหรือระหว่างบทโน้ตบุ๊ค	*		*				*		
มีการบูรณาการเนื้อหาได้อย่างเหมาะสม		*							
มีการแสดงการคำนวณโดยวิธีการทางสถิติ	*								
มีการเชื่อมโยงความรู้จากผลงานวิจัยดั้งเดิมกับความรู้อื่นๆ จากการวิจัย		*							
มีการอ้างอิงข้ามตอนให้ผู้อ่านติดตามศึกษาได้		*							
มีการเสนอข้อมูลเชิงลึกหรือรายละเอียดเด่นชัด		*							
5. ความครอบคลุมของเนื้อหา									
มีเนื้อหาสอดคล้องกับจุดหมายการสอน	*								
เนื้อหาครอบคลุมครบถ้วนสมบูรณ์	*	*	*		*				*
เนื้อหาในตำราเป็นเรื่องที่ผู้สอนสามารถสอนได้อย่างดี		*							
ตำรามีบทโน้ตบุ๊คตามที่ผู้เรียนคาดหวัง							*		
หัวข้อมีความหลากหลายกว้างขวาง	*	*			*				
หัวข้อมีสาระที่เหมาะสม	*	*							
เนื้อหามีรายละเอียดลึกซึ้ง		*			*				*
การให้เหตุผลเชิงนิรนัยและอุปนัยมีสัดส่วนที่เหมาะสม		*							
สารสนเทศมีความทันสมัย		*					*		
มีการเสนอสารสนเทศเก่าเพื่อบอกประวัติความเป็นมาเท่านั้น		*							
มีการแนะนำวิธีการใช้โปรแกรมสำเร็จรูปทางสถิติที่ทันสมัย					*				
มีการแนะนำการเลือกใช้โปรแกรมคอมพิวเตอร์ที่เหมาะสม					*				
เนื้อหามีส่วนเป็นนวัตกรรมใหม่ๆ						*			
เนื้อหาตำรามีคุณภาพดี							*		
6. ความถูกต้องของเนื้อหา									
คำอธิบายความเป็นมาถูกต้อง		*							

ตารางที่ 1 (ต่อ)

ประเด็นการประเมิน	Har	Re	Cobb	H&B	B&S	Sim	B&L	Jac	H
เนื้อหาปริมาณความกลาดเคลื่อนจากข้อเท็จจริงน้อย		*							
มโนทัศน์มีความถูกต้องและเป็นปรนัย						*			
ตัวอย่างถูกต้อง							*		
7. ตารางที่ทำให้เข้าใจดีขึ้น									
มีการอธิบายที่ชัดเจนเหมาะสม	*	*							
มีการอธิบายการพิสูจน์ที่เหมาะสม	*								
มีการอธิบายมโนทัศน์ชัดเจน			*				*		
มีการเน้นหลักพีชคณิตในการอธิบายความเกี่ยวข้องกับมโนทัศน์			*						
มีการเสนอสูตรใช้หลักพีชคณิตเมทริกซ์ที่เข้าใจง่าย				*					
มีการอธิบายแผนผังหรือโครงสร้าง	*								
มีการเสนอด้วยกราฟเส้น แผนภูมิที่เหมาะสม							*		
มีการอธิบายจากการเสนอด้วยกราฟ	*								
มีการอธิบายเชิงเปรียบเทียบ	*								
มีการอธิบายตัวอย่างที่สมบูรณ์เหมาะสม	*	*							
มีตัวอย่างที่เข้าใจง่าย							*		
การมีตัวอย่างที่ชัดเจน		*					*		
มีการเสนอหลักการวิเคราะห์ข้อมูลด้วยคอมพิวเตอร์					*				
มีการอธิบายขั้นตอนการหาค่าสถิติโดยใช้สูตรแต่ใช้ภาษาง่าย							*		
มีการอภิปรายการเลือกใช้สถิติที่เหมาะสม				*					
8. ลักษณะของแบบฝึกหัด									
โจทย์แบบฝึกหัดได้จากงานวิจัยจริง	*		*						
โจทย์สมมุติตามสถานการณ์จริง	*								
มีโจทย์ที่ทดสอบความสามารถ	*		*						
มีโจทย์น่าสนใจ	*								
มีโจทย์ยากมาก	*								
มีโจทย์ให้คำนวณ				*		*			
มีโจทย์ที่ช่วยให้ทักษะทางสถิติดีขึ้น	*						*		
ไม่มีโจทย์แบบฝึกหัด	*								
แบบฝึกหัดมีการจัดเรียงลำดับเป็นขั้นตอน		*							

ตารางที่ 1 (ต่อ)

ประเด็นการประเมิน	Har	Re	Cobb	H&B	B&S	Sim	B&L	Jac	Hol
แบบฝึกหัดมีความยากง่ายเหมาะสม		*							
แบบฝึกหัดมีความครอบคลุมการแปลความหมาย				*			*		
แบบฝึกหัดมีการเชื่อมโยงกับการบรรยาย					*				
แบบฝึกหัดให้ใช้ความคิดมากกว่าความจำ			*	*			*		
แบบฝึกหัดในตำรามีความเหมาะสม		*							
แบบฝึกหัดมีค่าเฉลย		*		*		*	*		
9. ส่วนเสริมของตำรา									
ตำรามีสารบัญเนื้อหาที่ถูกต้อง		*					*		
มีการเสนอเนื้อหาในเชิงวิทยาศาสตร์		*							
มีการอธิบายคำศัพท์		*					*		
มีคำสำคัญ สัญลักษณ์และคำนิยามชัดเจน		*					*		
มีการประยุกต์ทฤษฎีอย่างมีระบบ	*	*							
ตำรามีการสรุปท้ายบท		*							
ตำรามีเชิงอธิบายการหาค่าสถิติด้วยวิธีอื่นๆ						*			
ตำรามีแหล่งอ้างอิงที่ถูกต้อง		*							*
ตำรามีแหล่งอ้างอิงที่เพียงพอ		*							
ตำรามีภาคผนวกที่อำนวยความสะดวกต่อการศึกษาค้นคว้าด้วยตนเอง		*							
ตำรามีดัชนีที่สมบูรณ์		*					*		*
10. การให้คำแนะนำในการสอน									
มีคู่มือการสอนของครู		*							*
ในคู่มือครูมีแบบฝึกหัดเพิ่มเติมอย่างเหมาะสม		*							
มีคู่มือแนะนำการเรียน							*		
การมีแบบสอบถามวัดความรู้พื้นฐานก่อนเรียน						*			
การเรียงลำดับบทเรียนในทัศน์เหมาะสมกับแผนการสอน							*		
มีคำถามสำหรับการสอนที่เหมาะสมกับการตรวจด้วยคอมพิวเตอร์		*					*		
11. ลักษณะผู้แต่งตำรา									
ผู้แต่งมีคุณวุฒิที่น่าเชื่อถือ							*		

หมายเหตุ Har = Harwell, M. R. และคณะ ; Re = Redei, G. P. ; Cobb = Cobb, G. W. ;

H & B = Huberty, C. J. และ Barton, R. M. ; Sim = Simon, G. A. ;

B & S = Belli, G. M. และ Seaver, W. L. ; B & L = Berry, K. และ Lee, D. ;

Jac = Jacobson, M. D. ; Hol = Hollabaugh, M.

เหมาะสม (Redei, G. P. , 1984) และควมมีคุณภาพของตำราที่ดี (Berry, K และ Lee, D. , 1982)

2. รูปแบบการเขียน การประเมินในด้านนี้ มีประเด็นการประเมิน 10 ประเด็น คือ วิธีการเขียนถูกต้องตามหลักภาษา (Redei, G. P. , 1984) การใช้ถ้อยคำรัดกุมไม่ฟุ่มเฟือย (Redei, G. P. , 1984) การใช้ภาษาที่ง่าย (Redei, G. P. , 1984) การเขียนไม่เป็นระบบ ไม่มีระบบในการเขียน (Redei, G. P. , 1984) มีการเขียนเหมาะสมกับระดับความรู้ของผู้เรียน (Redei, G. P. , 1984) วิธีการเขียนอ่านแล้วเข้าใจง่าย (Redei, G. P. , 1984 ; Belli, G. M. และ Seaver, W. L. , 1989) มีการเขียนไม่เข้าประเด็นสำคัญของเรื่อง (Redei, G. P. , 1984) มีการเขียนที่ใช้ภาษาสละสลวย (Redei, G. P. , 1984) มีการเขียนเชิงอุปมา-อุปมัย (Redei, G. P. , 1984) มีการเขียนแบบบรรยาย (Redei, G. P. , 1984)

3. ความน่าอ่านของตำรา (readability) ประเด็นการประเมินในด้านนี้ ประกอบด้วย 11 ประเด็น คือ ระดับความยากของตำรา (Belli, G. M. และ Seaver, W. L. , 1989 ; Huberty, C. J. และ Barton, R. M. , 1990 ; Harwell, M. R. และคณะ , 1996) ระดับความยากในการอ่านตำรา (Redei, G. P. , 1984 ; Harwell, M. R. และคณะ , 1996) ระดับความลึกซึ้งทางวิชาการ (Cobb, G. W. , 1987) ระดับความยากในการคำนวณ (Hollabaugh, M. , 1989 ; Harwell, M. R. และคณะ , 1996) ระดับการใช้คำศัพท์ที่ยาก (Hollabaugh, M. , 1989) ดัชนีความยากในการอ่าน / ดัชนีความน่าอ่าน (Jacobson, M. D. , 1965 ; Hollabaugh, M. , 1989 ; Huberty, C. J. และ Barton, R. M. , 1990 ; Harwell, M. R. และคณะ , 1996)

4. วิธีการนำเสนอเนื้อหา การประเมินด้านวิธีการนำเสนอเนื้อหา นี้ ได้เสนอ 11 ประเด็น คือ วิธีการนำเสนอเนื้อหาเรียงจากง่ายไปยาก (Harwell, M. R. และคณะ , 1996) การเรียงลำดับหัวข้อสอดคล้องกับหัวข้อการสอน (Simon, G. A. , 1979 ; Harwell, M. R. และคณะ , 1996) การเสนอสาระมีความชัดเจน (Redei, G. P. , 1984) มีคำอธิบายมโนทัศน์ให้ผู้อ่านเข้าใจง่าย (Berry, K และ Lee, D. , 1982) การเสนอมโนทัศน์มีความกระชับไม่สับสน (Berry, K และ Lee, D. , 1982) มีการเชื่อมโยงเนื้อหาระหว่างหัวข้อหรือระหว่างมโนทัศน์ (Berry, K และ Lee, D. , 1982 ; Cobb, G. W. , 1987 ; Harwell, M. R. และคณะ , 1996) มีการบูรณาการเนื้อหาได้อย่างเหมาะสม (Redei, G. P. , 1984) มีการแสดงการคำนวณโดยวิธีทางสถิติ (Harwell, M. R. และคณะ , 1996) มีการเชื่อมโยงความรู้จากผลงานวิจัยดั้งเดิมกับความรู้ชายแดนจากการวิจัย (Redei, G. P. , 1984) มีการอ้างอิงข้ามตอนให้ผู้อ่านติดตามศึกษาได้ (Redei, G. P. , 1984) มีการเสนอข้อมูลเชิงลึกหรือรายละเอียดเด่นชัด (Redei, G. P. , 1984)

5. ความครอบคลุมของเนื้อหา การประเมินความครอบคลุมของเนื้อหาได้ใช้ประเด็นการประเมิน 14 ประเด็น คือ มีเนื้อหาสอดคล้องกับจุดหมายการสอน (Harwell, M. R. และคณะ , 1996) เนื้อหามีหัวข้อครบถ้วนสมบูรณ์ (Redei, G. P. , 1984 ; Cobb, G. W. , 1987 ; Belli, G. M. และ Seaver, W. L. , 1989 ; Hollabaugh, M. , 1989 ; Harwell, M. R. และคณะ , 1996) เนื้อหาในตำราเป็นเรื่องที่ผู้สอนสามารถสอนได้อย่างดี (Redei, G. P. , 1984) ตำรามีมโนทัศน์ตามที่ผู้เรียนคาดหวัง (Berry, K และ Lee, D. , 1982) หัวข้อมีความหลากหลายกว้างขวาง (Redei, G. P. , 1984 ; Belli, G. M. และ Seaver, W. L. , 1989 ; Harwell, M. R. และคณะ , 1996) หัวข้อมีสาระที่เหมาะสม (Redei, G. P. , 1984 ; Harwell, M. R. และคณะ , 1996) เนื้อหา มีรายละเอียดลึกซึ้ง (Redei, G. P. , 1984 ; Belli, G. M. และ Seaver, W. L. , 1989 ; Hollabaugh, M. , 1989) การให้เหตุผลเชิงนิรนัย - อุปนัยมีสัดส่วนที่เหมาะสม (Redei, G. P. , 1984) สารสนเทศมีความทันสมัย (Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984) มีการเสนอสารสนเทศเก่าเพื่อบอกประวัติความเป็นมาเท่านั้น (Redei, G. P. , 1984) มีการแนะนำวิธีการใช้โปรแกรมสำเร็จรูปทางสถิติที่ทันสมัย (Belli, G. M. และ Seaver, W. L. , 1989) มีการแนะนำการเลือกใช้โปรแกรมคอมพิวเตอร์ที่เหมาะสม (Belli, G. M. และ Seaver, W. L. , 1989) เนื้อหามีส่วนเป็นนวัตกรรมใหม่ๆ (Simon, G. A. , 1979) เนื้อหาตำรามีคุณภาพดี (Berry, K และ Lee, D. , 1982)

6. ความถูกต้องของเนื้อหา การประเมินด้านนี้ มีประเด็นการประเมิน 4 ประเด็น คือ คำอธิบายความเป็นมาถูกต้อง (Redei, G. P. , 1984) เนื้อหามีปริมาณความคลาดเคลื่อนจากข้อเท็จจริงน้อย (Redei, G. P. , 1984) มโนทัศน์มีความถูกต้องและเป็นปรนัย (Simon, G. A. , 1979) ตัวอย่างถูกต้อง (Berry, K และ Lee, D. , 1982)

7. สาระที่ทำให้เข้าใจดีขึ้น นอกจากเนื้อหาสำคัญของตำราแล้ว ยังมีสาระประกอบของเนื้อหาเสริมให้ผู้อ่านเกิดความเข้าใจดีขึ้น ในประเด็นการประเมินด้านนี้มี 16 ประเด็น คือ มีการอธิบายที่ชัดเจนเหมาะสม (Redei, G. P. , 1984 ; Harwell, M. R. และคณะ , 1996) มีการอธิบายพิสูจน์ที่เหมาะสม (Harwell, M. R. และคณะ , 1996) มีการอธิบายมโนทัศน์ชัดเจน (Berry, K และ Lee, D. , 1982 ; Cobb, G. W. , 1987) มีการเน้นหลักพีชคณิตในการอธิบายความเกี่ยวข้องกับมโนทัศน์ (Cobb, G. W. , 1987) มีการเสนอสูตรใช้หลักพีชคณิตเมทริกซ์ที่เข้าใจง่าย (Huberty, C. J. และ Barton, R. M. , 1990) มีการอธิบายแผนผังหรือโครงสร้าง (Harwell, M. R. และคณะ , 1996) มีการเสนอด้วยกราฟเส้น แผนภูมิที่เหมาะสม (Berry, K และ Lee, D. , 1982) มีการอธิบายจากการเสนอกราฟ (Harwell, M. R. และคณะ , 1996) มีการอธิบายเชิงเปรียบเทียบ (Harwell, M. R. และคณะ , 1996) มีการอธิบายตัวอย่างที่สมบูรณ์

เหมาะสม (Redei, G. P. , 1984 ; Harwell, M. R. และคณะ , 1996) มีตัวอย่างที่เข้าใจง่าย (Berry, K และ Lee, D. , 1982) การมีตัวอย่างที่ชัดเจน (Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984) มีการเสนอผลการวิเคราะห์ข้อมูลด้วยคอมพิวเตอร์ (Belli, G. M. และ Seaver, W. L. , 1989) มีการอธิบายขั้นตอนการหาค่าสถิติโดยไม่ใช้สูตรแต่ใช้ภาษาเขียน (Berry, K และ Lee, D. , 1982 ; Hollabaugh, M. , 1989) มีการอภิปรายการเลือกใช้สถิติที่เหมาะสม (Huberty, C. J. และ Barton, R. M. , 1990)

8. ลักษณะของแบบฝึกหัด การประเมินคุณภาพของแบบฝึกหัดนี้ จำเป็นต้องพิจารณาถึงลักษณะของแบบฝึกหัด ซึ่งในประเด็นการประเมินนี้มี 15 ประเด็น ได้แก่ โจทย์แบบฝึกหัดได้จากงานวิจัยจริง (Cobb, G. W. , 1987 ; Harwell, M. R. และคณะ , 1996) โจทย์สมมุติตามสถานการณ์จริง (Harwell, M. R. และคณะ , 1996) มีโจทย์ท้าทายความสามารถ (Cobb, G. W. , 1987 ; Harwell, M. R. และคณะ , 1996) มีโจทย์น่าสนใจ (Harwell, M. R. และคณะ , 1996) มีโจทย์ยากมาก (Harwell, M. R. และคณะ , 1996) มีโจทย์ให้คำนวณ (Simon, G. A. , 1979 ; Huberty, C. J. และ Barton, R. M. , 1990) มีโจทย์ที่ช่วยให้มีทักษะทางสถิติดีขึ้น (Berry, K และ Lee, D. , 1982 ; Harwell, M. R. และคณะ , 1996) ไม่มีโจทย์แบบฝึกหัด (Harwell, M. R. และคณะ , 1996) แบบฝึกหัดมีการจัดเรียงลำดับเป็นขั้นตอน (Redei, G. P. , 1984) แบบฝึกหัดมีความยากง่ายเหมาะสม (Redei, G. P. , 1984) แบบฝึกหัดมีความครอบคลุมการแปลความหมาย (Berry, K และ Lee, D. , 1982 ; Huberty, C. J. และ Barton, R. M. , 1990) แบบฝึกหัดมีการเชื่อมโยงกับการบรรยาย (Belli, G. M. และ Seaver, W. L. , 1989) แบบฝึกหัดให้ใช้ความคิดมากกว่าความจำ (Berry, K และ Lee, D. , 1982 ; Cobb, G. W. , 1987 ; Huberty, C. J. และ Barton, R. M. , 1990) แบบฝึกหัดในคำรามีความเหมาะสม (Redei, G. P. , 1984) แบบฝึกหัดมีค่าเฉลย (Simon, G. A. , 1979 ; Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984 ; Huberty, C. J. และ Barton, R. M. , 1990)

9. ส่วนเสริมของคำรามี คุณภาพของเนื้อหาของคำรามีแต่ละเล่มมีความสำคัญแล้ว แต่ยังมีส่วนประกอบภายในเล่มที่จะช่วยเสริมให้คำรามีเล่มนั้นมีความสมบูรณ์มากขึ้น ประเด็นที่ใช้สำหรับการประเมินด้านส่วนเสริมของคำรามี มี 11 ประเด็น คือ คำรามีสารบัญเนื้อหาที่ถูกต้อง (Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984) มีการเสนอเนื้อหาในเชิงวิทยาศาสตร์ (Redei, G. P. , 1984) มีการอธิบายคำศัพท์ (Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984) มีคำสำคัญ สัญลักษณ์ และคำนิยามชัดเจน (Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984) มีการประยุกต์ทฤษฎีอย่างมีระบบ (Redei, G. P. , 1984 ; Harwell, M. R. และคณะ , 1996) คำรามีมีการสรุปท้ายบท (Redei, G. P. , 1984) คำรามีเชิงอรรถอธิบายการหาค่าสถิติด้วยวิธีอื่นๆ

(Simon, G. A. , 1979) คำรามมีแหล่งอ้างอิงที่ถูกต้อง (Redei, G. P. , 1984) คำรามมีแหล่งอ้างอิงที่เพียงพอ (Redei, G. P. , 1984) คำรามมีภาคผนวกที่อำนวยความสะดวกต่อการศึกษาด้วยตนเอง (Redei, G. P. , 1984) คำรามมีดัชนีที่สมบูรณ์ (Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984)

10. การให้คำแนะนำในการสอน ในส่วนการประเมินการให้คำแนะนำในการสอนนี้ มีประเด็นการประเมิน 6 ประเด็น คือ มีคู่มือการสอนของครู (Redei, G. P. , 1984) ในคู่มือครูมีแบบฝึกหัดเพิ่มเติมอย่างเหมาะสม (Redei, G. P. , 1984) มีคู่มือแนะนำการเรียน (Berry, K และ Lee, D. , 1982) การมีแบบสอบวัดความรู้พื้นฐานก่อนเรียน (Simon, G. A. , 1979) การเรียงลำดับบทโน้ตที่ตรงกับแผนการสอน (Berry, K และ Lee, D. , 1982) มีคำถามสำหรับการสอบที่เหมาะสมกับการตรวจด้วยคอมพิวเตอร์ (Berry, K และ Lee, D. , 1982 ; Redei, G. P. , 1984)

11. ลักษณะผู้แต่งตำรา ในการเขียนตำราเพื่อการศึกษา ผู้แต่งเป็นส่วนหนึ่งที่มีความสำคัญต่องานเขียนตำราที่ดีมีคุณภาพและเป็นประโยชน์ต่อผู้อ่านหรือผู้เรียนในการใช้ประกอบการเรียนการสอน ผู้แต่งต้องมีความรู้และเชี่ยวชาญในสาขานั้นเป็นอย่างดี การประเมินด้านลักษณะผู้แต่งตำรามีประเด็นการประเมิน คือ ผู้แต่งมีคุณวุฒิที่น่าเชื่อถือ (Berry, K และ Lee, D. , 1982)

เนื่องจากประเด็นการประเมินด้านดัชนีความน่าอ่าน มีวิธีการคำนวณที่ซับซ้อนแตกต่างจากประเด็นอื่น เพราะในการประเมินต้องมีการรวบรวมข้อมูลมาพัฒนาดัชนี ซึ่งมีวิธีการพัฒนาแตกต่างกันหลายแบบ ดังนั้นผู้วิจัยจึงนำเสนอสาระเกี่ยวกับการพัฒนาดัชนีความน่าอ่าน โดยนำเสนอดัชนี 3 แบบที่ใช้ในการประเมินตำราสถิติ ได้แก่ ดัชนีของ Jacobson ดัชนีของ Berry และ Lee ดัชนีของ Hollabaugh ดังรายละเอียดต่อไปนี้

ดัชนีตัวที่ 1. ดัชนีของ Jacobson

การพัฒนาดัชนีความยากในการอ่านตามวิธีของ Jacobson, M. D. (1965) เป็นการวัดระดับความยากในการอ่านตำราฟิสิกส์และเคมี Jacobson ได้ใช้แบบสอบขีดเส้นใต้ (underlining test) ที่มีตัวแปรรวม 5 ตัว ได้แก่ จำนวนคำโดยเฉลี่ยต่ออนุประโยค จำนวนคำศัพท์คณิตศาสตร์ที่ยากในข้อความที่กำหนด จำนวนคำศัพท์ที่มีระดับความยากสูงกว่าระดับความยากของคำศัพท์ 6,000 คำในศัพท์ของ Thorndike รวมทั้งหมด 20,000 คำ จำนวนคำศัพท์วิทยาศาสตร์ที่ไม่อยู่ในรายการคำศัพท์ 1,828 คำของ Power และอัตราส่วนของจำนวนตัวอักษรโดยเฉลี่ยต่อคำกับจำนวนคำในข้อความที่กำหนดให้

วิธีการหาดัชนีความยากในการอ่าน ตามสูตรของ Jacobson (Jacobson's formula) มีวิธีดำเนินการสร้างแบบสอบขีดเส้นได้ และการพัฒนาดัชนีความยากในการอ่าน ดังนี้คือ

ก. แบ่งเนื้อหาในตำราแต่ละเล่มเป็น 10 ส่วน ในแต่ละส่วนประกอบด้วยเนื้อหาสาระย่อยละ 10 ของตำรา

ข. ตัดเนื้อหาแต่ละส่วนให้ได้ข้อความมีความยาวประมาณ 200 คำ ได้ข้อความที่จะใช้ในการประเมินเล่มละ 10 ข้อความ นำมาจัดทำเป็นหนังสือเล่มเล็ก (booklets) พร้อมทั้งคำชี้แจงสำหรับผู้ประเมิน

ค. นำหนังสือเล่มเล็กที่มี 10 ข้อความให้ผู้ประเมินขีดเส้นได้คำที่ยาก และผู้วิจัยนับจำนวนคำที่ขีดเส้นได้ดังกล่าว คำที่ได้จะเป็นค่าของตัวแปรตามในการพัฒนาดัชนีความยากในการอ่าน

ง. ผู้วิจัยหาค่าของตัวแปรอิสระรวม 5 ตัวสำหรับการพัฒนาดัชนีความยากในการอ่าน ดังนี้

- 1) จำนวนคำโดยเฉลี่ยต่ออนุประโยค
- 2) จำนวนคำศัพท์ทางคณิตศาสตร์ที่ยาก ในข้อความที่กำหนด
- 3) จำนวนคำศัพท์ที่ระดับความยากสูงกว่าระดับความยากของคำศัพท์ 6,000 คำในศัพท์ของThorndike รวมทั้งหมด 20,000 คำ
- 4) จำนวนคำศัพท์วิทยาศาสตร์ที่ไม่อยู่ในรายการคำศัพท์ 1,828 คำของ Power
- 5) อัตราส่วนของจำนวนตัวอักษรโดยเฉลี่ยต่อคำ กับจำนวนคำในข้อความที่กำหนดให้

จ. นำตัวแปรตามและตัวแปรอิสระมาวิเคราะห์สมการถดถอยพหุคูณ จะได้สมการพยากรณ์ดัชนีความยากในการอ่านจากตัวแปรอิสระทั้ง 5 ตัว ใช้ประโยชน์ต่อไป

ดัชนีตัวที่ 2 ดัชนีของ Berry และ Lee

หลักในการคำนวณดัชนีความนำอ่านของ Berry, K. และ Lee, D. (1982) เป็นหลักการคำนวณดัชนีความนำอ่านของ Fry, E. (1977) อ้างถึงใน Berry และ Lee, (1982) วิธีการนี้ใช้ความสัมพันธ์ของจำนวนประโยคในข้อความที่มีความยาว 100 คำ กับจำนวนพยางค์โดยเฉลี่ยในข้อความที่มีความยาว 100 คำ โดยมีขั้นตอนการพัฒนาดังนี้

ก. เลือกข้อความโดยการสุ่มข้อความ (passage) ในตำรา 3 ข้อความ ให้มีความยาวข้อความละ 100 คำ ทั้งนี้การนับจำนวนคำในแต่ละข้อความ นับรวมทั้ง วิสามานยนาม ตัวย่อ และตัวเลข

ข. นับจำนวนประโยคในข้อความที่มีความยาว 100 คำ ที่สุ่มเลือกไว้ โดยนับส่วนของประโยคสุดท้ายของข้อความเป็นตัวเลขทศนิยม 1 ตำแหน่ง

ก. นับจำนวนพยางค์ทั้งหมดในข้อความที่มีความยาว 100 คำ ด้วยการขีดเส้นใต้ทำเครื่องหมายพยางค์การออกเสียงแต่ละพยางค์ แล้วนับจำนวนเครื่องหมายที่ขีดเส้นใต้นั้นแล้วบวกด้วยจำนวน 100

ง. สร้างกราฟแสดงความสัมพันธ์ระหว่างจำนวนประโยคโดยเฉลี่ยต่อข้อความที่มีความยาว 100 คำ และจำนวนพยางค์เฉลี่ยในข้อความนั้นแล้วลงจุดเครื่องหมายแสดงตำแหน่งที่เส้นกราฟตัดกันอันเป็นตำแหน่งที่บอกระดับความน่าอ่านของข้อความ

ดัชนีตัวที่ 3. ดัชนีหมอกหรือดัชนีความไม่ชัดเจนของ Gunning ในงานวิจัยของ Hollabaugh

ในการประเมินตำราดาราศาสตร์ ระดับอุดมศึกษา Hollabaugh, M. (1989) ใช้การประเมินโดยวัดดัชนีความน่าอ่านตามดัชนีหมอกหรือดัชนีความไม่ชัดเจนของ Gunning (Gunning Fog Index) ซึ่งเป็นดัชนีบอกความยากของข้อความ Hollabaugh, M. ใช้ดัชนีหมอกหรือดัชนีความไม่ชัดเจนในการเปรียบเทียบความน่าอ่านของตำรา 2 เล่ม โดยคำนวณดัชนีหมอกหรือดัชนีความไม่ชัดเจนของข้อความที่มีสระแบบเดียวกัน การเปรียบเทียบดังกล่าวทำได้ เนื่องจากดัชนีหมอกหรือดัชนีความไม่ชัดเจนใช้หลักการนับจำนวนคำที่มี 3 พยางค์ขึ้นไปเป็นหลักในการคำนวณ ซึ่งมีขั้นตอนการดำเนินการดังนี้

ก. คัดเลือกข้อความจากตำรา 1 ตอน ให้มีความยาวอย่างน้อย 100 คำ

ข. นับจำนวนประโยคในข้อความที่มีความยาว 100 คำนั้น

ค. คำนวณหาค่าเฉลี่ยของจำนวนคำในประโยค ซึ่งคำนวณได้จากจำนวนคำในข้อความหารด้วยจำนวนประโยค

ง. นับจำนวนคำที่มี 3 พยางค์ขึ้นไปในข้อความที่มีความยาว 100 คำ นั้น

จ. รวมค่าเฉลี่ยจำนวนคำในประโยคจากข้อ ค กับจำนวนคำที่มี 3 พยางค์จากข้อ ง

ฉ. นับผลลัพธ์จากข้อ จ คูณด้วย 0.4 และมีจุดทศนิยมให้เป็นเลขจำนวนเต็ม ผลลัพธ์ที่ได้เป็นดัชนีความน่าอ่านของตำรา ถ้าดัชนีมีค่าสูงแสดงว่าตำรามีระดับความยากมาก ถ้าดัชนีมีค่าต่ำแสดงว่าตำรามีระดับที่ง่าย

จากการศึกษาวิธีการพัฒนาดัชนีค่าความน่าอ่าน/ดัชนีความยากในการอ่านที่เสนอมาย่างค้นสรุปได้ว่า การพัฒนาดัชนีความน่าอ่าน/ดัชนีความยากในการอ่านทำได้ 3 รูปแบบคือ ดัชนีของ Jacobson ซึ่งคำนวณจากรูปแบบสมการพยางค์ (Jacobson, M. D. , 1965) ดัชนีของ Barry

และ Lee ซึ่งประมาณค่าจากกราฟ (Berry, K. และ Lee, D. , 1982) และดัชนีหมอก หรือดัชนีความไม่ชัดเจนของ Gunning ในงานวิจัยของ Hollabaugh (Hollabaugh, M. , 1989) ผู้วิจัยพิจารณาเห็นว่าดัชนีความน่าอ่านในงานวิจัยของ Hollabaugh เป็นดัชนีที่บอกระดับความยากของตำราที่มีการคำนวณง่ายสะดวกและตีความหมายได้ง่าย จึงพิจารณาเลือกใช้ดัชนีความน่าอ่านที่ใช้ในงานวิจัยของ Hollabaugh ในการวิจัยครั้งนี้

3.3 วิธีดำเนินการประเมินตำราสถิติ

จากการวิเคราะห์และการปรัศนงานวิจัยที่เป็นการประเมินหรือการวิเคราะห์เชิงตำราสถิติตามที่เสนอในหัวข้อ 3.2 นั้น ผู้วิจัยพบว่าการประเมินทำได้หลายวิธี ซึ่งนำเสนอวิธีการประเมินตำราสถิติวิธีการประเมินแต่ละวิธีโดยนำเสนอ 5 ประเด็น คือ กลุ่มตัวอย่างหรือผู้ประเมิน เครื่องมือที่ใช้ในการประเมิน วิธีการเก็บรวบรวมข้อมูล วิธีการวิเคราะห์ข้อมูล และผลการประเมิน โดยนำเสนอแยกตามประเภทของผลงานเป็น 2 กลุ่ม กลุ่มแรกเป็นวิธีการประเมินตำราสถิติจากผลงานวิจัยเชิงประเมิน 5 เรื่องที่สำคัญ ได้แก่ งานวิจัยของ Harwell, M. R. และคณะ (1996) ; Huberty, C. J. และ Barton, R. M. (1990); Cobb, G. W. (1987) และ Berry, K. และ Lee, D. (1982) กลุ่มที่ 2 เป็นวิธีการประเมินที่มีการพัฒนาดัชนีจากผลงานวิจัยเชิงประเมิน 2 เรื่อง ได้แก่ งานวิจัยของ Jacobson, M. D. (1965) และ Hollabaugh, M. (1989) ดังรายละเอียดต่อไปนี้

เรื่องที่ 1 Harwell, M. R. (1996) ได้ประเมินคุณภาพตำราสถิติเบื้องต้น จำนวน 6 เล่ม โดยใช้ผู้ประเมิน แบ่งเป็น 3 กลุ่ม คือ กลุ่มนักศึกษาที่เรียนวิชาสถิติจำนวน 253 คน กลุ่ม ผู้สอนวิชาสถิติจำนวน 8 คน และกลุ่มผู้เชี่ยวชาญจำนวน 6 คน เครื่องมือที่ใช้ในการประเมิน เป็นแบบประเมิน 3 ชุด คือ แบบประเมินสำหรับผู้เรียน แบบประเมินสำหรับผู้สอน และแบบประเมินสำหรับผู้เชี่ยวชาญ ดังรายละเอียดที่จะกล่าวต่อไป คือ

1. แบบประเมินสำหรับผู้เรียน แบบประเมินแบ่งออกเป็น 2 ส่วน คือ ส่วนที่ 1 เป็นแบบสอบถามภูมิหลังการเรียนคณิตศาสตร์และสถิติ ส่วนที่ 2 เป็นข้อคำถามประเมินตำราสถิติด้านการอธิบาย ลีลาการเขียน การเน้นเนื้อหาในภาพรวม ลักษณะการประเมินเป็นแบบประเมินประเภทมาตรประเมินค่า (rating scale)

2. แบบประเมินสำหรับผู้สอน แบบประเมินแบ่งเป็น 2 ส่วน ส่วนแรกเป็นแบบสอบถามเกี่ยวกับลักษณะผู้ตอบและภูมิหลังการสอนวิชาสถิติ ส่วนที่ 2 เป็นการประเมินตำราสถิติเช่นเดียวกับแบบประเมินสำหรับผู้เรียน แต่แบบประเมินของผู้สอนจะมีข้อคำถามเพิ่มมากกว่าแบบประเมิน

ของผู้เรียน ตรงส่วนที่ถามเกี่ยวกับโปรแกรมวิเคราะห์ข้อมูลทางสถิติ และมีคำถามปลายเปิดเกี่ยวกับความคิดเห็นการใช้โปรแกรม จุดติจุดด้อยของตำรา

3. แบบประเมินสำหรับผู้เชี่ยวชาญ เป็นแบบประเมินด้านคุณภาพและความถูกต้องของเนื้อหาทางสถิติในตำรา 4 หัวข้อ แบ่งเป็นหัวข้อด้านสังกัดการทดสอบสมมุติฐาน 2 หัวข้อคือ ความคลาดเคลื่อนประเภทที่ I และ II (type I and type II error) และสังกัดเกี่ยวกับสถิติอนุมาน 2 หัวข้อ คือ การใช้การทดสอบไค-สแควร์ (χ^2 - test) และปฏิสัมพันธ์ในการวิเคราะห์ความแปรปรวนแบบสองทาง ลักษณะเครื่องมือการประเมินเป็นมาตราประเมินค่าแบบลิเคอร์ท (Likert scale) ที่มีน้ำหนักคะแนนเป็น 1 - 5 (1 = มีค่าไม่เห็นด้วย, 5 = มีค่าเห็นด้วยอย่างยิ่ง) และคำถามปลายเปิด

แบบประเมินตำราทั้ง 3 ชุด มีความตรงเชิงเนื้อหา (content validity) และมีการศึกษานำร่องของเครื่องมือเพื่อปรับปรุงแก้ไขก่อนนำไปใช้จริงกับกลุ่มตัวอย่าง การเก็บรวบรวมข้อมูลจริงใช้ระยะเวลา 1 เดือนสุดท้ายของการเรียนในเทอมการศึกษาต้น และระยะเวลา 2 - 4 สัปดาห์แรกของการเรียนในภาคการศึกษาที่ 2 โดยใช้ระยะเวลา 15 - 45 นาทีในการตอบแบบประเมิน

วิธีการวิเคราะห์ข้อมูล แบ่งเป็น 2 แบบ การวิเคราะห์ข้อมูลที่เป็นภูมิหลังของผู้ประเมินใช้สถิติบรรยาย เช่น ร้อยละ การวิเคราะห์ข้อมูลที่เป็นผลการประเมินจากแบบประเมินตำราสถิติใช้การคำนวณค่าสถิติ เช่น มัชฌิมเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson's product moment correlation coefficient) และการวิเคราะห์ความแปรปรวนเพื่อเปรียบเทียบค่าเฉลี่ยของกลุ่มประชากร 2 กลุ่ม

ผลของการวิเคราะห์ข้อมูล พบว่าแบบประเมินประเภทมาตราประเมินค่าแบบลิเคอร์ทที่สร้างขึ้นใช้กับผู้ประเมินที่เป็นผู้เรียน ผู้สอน และผู้เชี่ยวชาญ ให้ผลการประเมินตำราสถิติสอดคล้องกัน นอกจากนี้ยังจำแนกความแตกต่างของคุณภาพตำราแต่ละเล่มได้ว่า ตำราเล่มไหนมีความน่าอ่าน ตำราเล่มใดที่อ่านง่าย และการคำนวณอยู่ระดับใด

เรื่องที่ 2. Huberty, C. J. และ Barton, R. M. (1990) ได้ประเมินคุณภาพของตำรา การวิเคราะห์ตัวแปรพหุนาม จำนวน 15 เล่ม โดยมีจุดมุ่งหมายเพื่อเสนอผลการประเมินในรูปการปริทัศน์ตำรา เครื่องมือสำหรับการประเมิน เป็นแบบประเมินประเภทมาตราประเมินค่า (rating scale) 5 ระดับ ประกอบด้วยประเด็นการประเมิน 4 ประเด็น คือ ความครอบคลุมเนื้อหา การแนะนำให้ลงมือทำ ความน่าอ่าน และคุณภาพของแบบฝึกหัด การวิเคราะห์ข้อมูล ใช้ค่าสถิติร้อยละ และเสนอข้อมูลด้านคุณภาพของตำรา โดยการเปรียบเทียบคะแนนรวมของคะแนนการประเมิน

ค่าคะแนนรวมที่มีค่าสูงจะเป็นหนังสือที่มีคุณภาพมากที่สุด

ผลการประเมินให้ภาพรวม คุณภาพของตำราในด้าน ความครอบคลุมของเนื้อหา การแนะนำให้ลงมือทำ ความน่าอ่าน และคุณภาพของแบบฝึกหัดของตำราทั้ง 15 เล่ม ให้สารสนเทศที่ช่วยในการคัดเลือกตำราได้

เรื่องที่ 3. Cobb, G. W. (1987) ได้เสนอกรอบแนวทางการประเมิน (framework for evaluation) เพื่อประเมินคุณภาพของตำราสถิติ และทำการเปรียบเทียบผลการประเมินตำราสถิติเบื้องต้นจำนวน 16 เล่ม ใน 3 ประเด็น คือ ระดับวิชาการและคุณภาพของการอธิบายสาระ ความครอบคลุมของหัวข้อ และคุณภาพของแบบฝึกหัด

วิธีการประเมินในประเด็นระดับวิชาการและคุณภาพการอธิบายสาระ ใช้การเลือกสาระในตำรา 4 เรื่อง คือ ค่าเฉลี่ยกลุ่มตัวอย่าง การแจกแจงปกติ และการแจกแจงของค่าสถิติจากกลุ่มตัวอย่าง แล้วประเมินว่าผู้แต่งใช้สูตรและศัพท์สถิติในการเสนอเนื้อหาสาระมากน้อยเพียงใด วิธีการประเมินในประเด็นความครอบคลุมของหัวข้อ ใช้วิธีตรวจสอบความสมบูรณ์ของสาระใน 6 หัวข้อ คือ การวิเคราะห์การถดถอย การวิเคราะห์ความแปรปรวน การวิเคราะห์เบื้องต้น การใช้คอมพิวเตอร์ ความน่าจะเป็น และสถิตินัยพารามตริก สำหรับวิธีการประเมินด้านคุณภาพของแบบฝึกหัด พิจารณาแยก 3 ประเด็น คือ ข้อมูลจริงหรือข้อมูลจำลอง การทำแบบฝึกหัดช่วยให้เกิดความเข้าใจหรือไม่ และการทำแบบฝึกหัดทำให้ผู้อ่านได้คิดเป็นสัดส่วนเท่าไร

ผลการประเมิน Cobb, G. W. ได้นำเสนอระดับคุณภาพของตำราแยกเป็น 9 ระดับ ตั้งแต่ระดับ 1 เป็นตำราที่มีความง่ายชัดเจน ไปจนถึงระดับ 9 จัดเป็นตำราที่มีความยากสูงสุด น่าสนใจ และทำให้เกิดการเรียนรู้ได้มาก

เรื่องที่ 4. Berry, K. และ Lee, D. (1982) ได้ประเมินตำราเคมีจำนวน 23 เล่ม โดยพิจารณาจากความยากของตำรา ในการวิจัยของเขา Berry, K. และ Lee, D. ใช้แนวทางการคำนวณค่าดัชนีความน่าอ่านของ Fry, E. และนำผลการคำนวณมาเปรียบเทียบกับดัชนีความน่าอ่านที่ได้จากงานวิจัยของ Flesch นอกจากนี้ Berry, K. และ Lee, D. ยังประเมินคุณภาพของตำราในด้านอื่นๆอีก 9 ด้าน โดยใช้แบบประเมินชนิดเลือกตอบ รวม 31 ข้อ เป็นเครื่องมือสำหรับการประเมิน แล้วนำผลการประเมินมาเปรียบเทียบกัน

ผลการประเมินพบว่าคุณภาพของคำราแตกต่างกัน แต่ Berry, K และ Lee, D มีได้ตัดสินใจว่า คำราเล่มใดดีกว่ากัน เพียงแต่ให้ข้อมูลสารสนเทศสำหรับนักวิชาการ นักศึกษาใช้ประกอบการตัดสินใจเลือกซื้อหรือเลือกใช้คำรา

เรื่องที่ 5 Jacobson, M. D. (1965) ได้ประเมินคำราฟิสิกส์และเคมี เพื่อคัดเลือกคำราสำหรับประกอบการเรียนการสอน โดยประเมินคำราฟิสิกส์จำนวน 13 เล่ม และคำราเคมีจำนวน 13 เล่ม ในการประเมินครั้งนี้ ผู้ประเมินได้แก่นักเรียนที่เรียนวิชาฟิสิกส์จากโรงเรียนของรัฐระดับชั้นมัธยมศึกษาตอนปลายมลรัฐมินเนโซต้า จำนวน 12 โรงเรียน โรงเรียนละ 16 คน และนักเรียนที่เรียนวิชาเคมีจำนวน 10 โรงเรียน โรงเรียนละ 16 คน โดยมีจุดมุ่งหมายสำคัญคือการประเมินความยากของคำรา

เรื่องมือที่ใช้ในการประเมินเป็นแบบสอบขีดเส้นใต้ ประกอบด้วยข้อความที่สุ่มมาจากคำรา และให้ผู้ประเมินขีดเส้นใต้คำศัพท์ที่ยาก อ่านไม่เข้าใจ จากนั้นผู้ประเมินรวบรวมข้อมูลผลการประเมินที่ได้จากผู้ประเมิน และข้อมูลที่ผู้ประเมินรวบรวมด้วยตนเองมากำหนดดัชนีความยาก และสร้างสมการพยากรณ์ดัชนีความยาก ผู้ประเมินเสนอการวิเคราะห์เป็นสถิติบรรยาย ได้แก่ ค่าเฉลี่ย ค่าความเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์สหสัมพันธ์ของ Pearson การวิเคราะห์การถดถอยพหุคูณ การวิเคราะห์ความแปรปรวน เปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มด้วยวิธีของ Scheffe

ผลการวิเคราะห์ข้อมูล พบว่าคุณภาพด้านความยาก คำราฟิสิกส์และคำราเคมีมีความแตกต่างกันตามโรงเรียน ตามนักเรียนผู้ประเมิน และตามหัวข้อเนื้อหาในคำราที่ประเมิน นอกจากนี้ยังได้สมการพยากรณ์ดัชนีความยาก ซึ่งจะนำไปใช้ในการประเมินคำราต่อไป

เรื่องที่ 6 Hollabaugh, M. (1989) ได้ประเมินคำราดาราศาสตร์จำนวน 5 เล่ม เพื่อคัดเลือกคำราสำหรับผู้เรียนและผู้สอนใช้ประกอบการเรียนการสอนวิชาดาราศาสตร์ การประเมินเป็นการพิจารณาเฉพาะประเด็นความยาก โดยใช้ดัชนีหมอกหรือดัชนีความไม่ชัดเจน เป็นตัวบ่งชี้ Hollabaugh ได้กำหนดเกณฑ์การประเมินไว้ว่า คำราที่มีค่าดัชนีหมอกหรือดัชนีความไม่ชัดเจนเป็น 15 หมายถึง คำราที่มีความยากเทียบเท่าระดับปริญญาตรี และคำราที่มีค่าดัชนีหมอกหรือดัชนีความไม่ชัดเจนตั้งแต่ 20 ขึ้นไป หมายถึง คำราที่มีความยากเทียบเท่าระดับบัณฑิตศึกษาถึงระดับปริญญาเอก

Hollabaugh, M. ได้คำนวณดัชนีหมอกหรือดัชนีความไม่ชัดเจน จากตำราดาราศาสตร์ทั้ง 5 เล่ม และนำไปเปรียบเทียบกัน โดยเปรียบเทียบเนื้อหาสาระแยกกันเป็น 5 ตอน และแสดงให้เห็นว่าตำราแต่ละเล่มมีค่าดัชนีหมอกหรือดัชนีความไม่ชัดเจนในแต่ละตอนแตกต่างกัน

จากการศึกษาเปรียบเทียบวิธีการประเมินตำราจากงานวิจัย 6 เรื่องที่กล่าวมาแล้วข้างต้น ผู้วิจัยสรุปวิธีการประเมินในหัวข้อ ประเภทของการประเมิน ประชากรและกลุ่มตัวอย่าง เครื่องมือ สำหรับการประเมินตำรา วิธีการเก็บรวบรวมข้อมูล และการวิเคราะห์ข้อมูล ดังรายละเอียดต่อไปนี้

ก. ประเภทของการประเมิน การประเมินตำราสถิติและตำราวิทยาศาสตร์ที่ผู้วิจัยศึกษาทั้งหมด เป็นการประเมินตำราเพื่อคัดเลือกตำราใช้ประกอบการเรียนการสอน

ข. ประชากรและกลุ่มตัวอย่างที่ใช้ในการประเมิน ประชากรและกลุ่มตัวอย่างประกอบด้วยบุคคล 3 กลุ่ม คือ ผู้สอนวิชาสถิติ ผู้เรียนวิชาสถิติ และผู้เชี่ยวชาญ ผู้สอนวิชาสถิติเป็นผู้มีความรู้ทางสาขาสถิติและประกอบอาชีพเป็นผู้สอนศาสตร์สาขานี้ อีกทั้งเป็นผู้คัดเลือกตำราสถิติใช้เป็นอุปกรณ์การสอนและเพื่อดำเนินการสอนให้ตรงตามวัตถุประสงค์ของหลักสูตร ผู้สอนจึงเป็นผู้มีบทบาทสำคัญในการประเมินตำรา (Huberty และ Barton , 1990) ส่วนผู้เรียน ที่ศึกษาสาขาสถิติและใช้ตำราสถิติประกอบการเรียนการสอนนั้น Hollabaugh, M. (1996) กล่าวว่า เป็นกลุ่มผู้มีความรู้ความสามารถเพียงพอที่จะประเมินตำราสถิติได้ทั้งในการประเมินด้านทัศนคติต่อตำราและความยากในการอ่านตำรา กลุ่มประชากรผู้ประเมินกลุ่มสุดท้าย คือ ผู้เชี่ยวชาญในสาขาสถิติสามารถตรวจสอบความตรงของเนื้อหา และความถูกต้องตามหลักวิชา รวมทั้งสามารถประเมินคุณภาพตำราได้ทุกประเด็น

ค. เครื่องมือสำหรับการประเมิน เครื่องมือที่ใช้ในการประเมินเป็นแบบประเมินที่ผู้วิจัยสร้างขึ้นเอง ในงานวิจัยทั้ง 6 เรื่อง จะมีแบบประเมินสำหรับผู้ประเมินแต่ละกลุ่ม ได้แก่ แบบประเมินสำหรับผู้เรียน แบบประเมินสำหรับผู้สอน และแบบประเมินสำหรับผู้เชี่ยวชาญ แบบประเมินทั้ง 3 แบบ ส่วนใหญ่เป็นแบบประเมินประเภทมาตราประเมินค่า (rating scale) แบบลิเคอร์ท (Likert) ที่ให้น้ำหนักคะแนนเป็น 1 - 5 (1 = น้อยที่สุด และ 5 = มากที่สุด) และมีแบบประเมินบางส่วนเป็นแบบสอบถามชนิดเลือกตอบและเติมค่าในช่องว่าง แบบประเมินแต่ละชุดใช้เวลาประมาณ 15 - 45 นาที

ง. การวิเคราะห์ข้อมูล การวิเคราะห์ข้อมูล เพื่อเสนอผลการประเมินบ่งบอกคุณภาพของตำรา ส่วนใหญ่ไม่มีการกำหนดเกณฑ์การประเมินแต่มีการอภิปรายผลโดยเทียบกับเกณฑ์ที่ผู้วิจัยกำหนดไว้ วิธีการวิเคราะห์ข้อมูลใช้วิธีแตกต่างกันหลายแบบสรุปได้ดังนี้

1. การวิเคราะห์ข้อมูลพื้นฐานโดยใช้ค่าสถิติบรรยาย เช่น ร้อยละ การแจกแจงความถี่ ค่ามัธยฐานเลขคณิต ค่าส่วนเบี่ยงเบนมาตรฐาน
2. การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน เพื่อแสดงความสัมพันธ์ระหว่างผลการประเมินตั้งแต่ 2 ประเด็นขึ้นไป
3. การวิเคราะห์ความแปรปรวน หรือการวิเคราะห์เปรียบเทียบ ค่าเฉลี่ยผลการประเมิน เพื่อเปรียบเทียบว่าผู้ประเมินตั้งแต่ 2 กลุ่มขึ้นไปมีความคิดเห็นว่าคุณภาพตำราแตกต่างกันอย่างไร
4. การพัฒนาดัชนี (index) หรือตัวบ่งชี้ (indicators) เพื่อสรุปผลการประเมินในประเด็นย่อยๆ ให้ได้ตัวบ่งชี้รวมที่บอกคุณภาพในภาพรวมของตำรา

โดยสรุปวิธีการประเมินที่ผู้วิจัยใช้ในการวิจัยครั้งนี้ เป็นวิธีการที่งานวิจัยส่วนใหญ่ใช้กัน คือ การประเมินโดยใช้ผู้สอนและนิสิตเป็นผู้ประเมิน เครื่องมือในการประเมินใช้มาตรฐานประเมินค่าแบบ Likert ซึ่งผู้วิจัยจะรวบรวมด้วยตนเอง จากนั้นจึงวิเคราะห์ข้อมูลโดยใช้สถิติบรรยายวิเคราะห์ข้อมูลพื้นฐาน ในการวิเคราะห์เปรียบเทียบค่าเฉลี่ย เพื่อเปรียบเทียบผลการประเมินจากผู้ประเมิน 2 กลุ่ม และการพัฒนาตัวบ่งชี้เพื่อสรุปผลการประเมิน

3.4 เกณฑ์การประเมินตำรา

ในกระบวนการประเมินหรือการวิเคราะห์เชิงประเมิน นอกจากนักประเมินจะต้องกำหนดจุดมุ่งหมายของการประเมิน ประเด็นการประเมิน วิธีดำเนินการประเมิน ดังที่ได้เสนอมาแล้วนั้น ยังมีส่วนต้องกำหนดอีกอย่างหนึ่ง ซึ่งเป็นส่วนที่มีความสำคัญต่อการพิจารณาตัดสินคุณค่าสิ่งที่จะประเมิน คือ เกณฑ์การประเมิน ในการเสนอรายงานเอกสารที่เกี่ยวข้องในหัวข้อนี้ ผู้วิจัยนำเสนอสาระเกี่ยวกับเกณฑ์การประเมิน เพื่อให้ทราบความหมายและวิธีการกำหนดเกณฑ์การประเมินดังนี้

การประเมินคุณค่าของสิ่งที่ต้องการประเมิน นักประเมินจำเป็นต้องมี มาตรฐาน (standard) หรือ เกณฑ์ (criteria) สำหรับการตรวจสอบคุณภาพของสิ่งที่ประเมินเพื่อประกอบการตัดสินใจ ในสาขาการประเมินมีผู้ให้ความหมายของเกณฑ์แตกต่างกันไป ดังที่จะกล่าวในลำดับต่อไป

พัทวิ จริตรธรรม (2539) ได้รายงานว่า เกณฑ์ในความหมายของราชบัณฑิตยสถาน หมายถึง หลักที่กำหนดไว้ในการวินิจฉัย และ Cizek, G. J. (1993) ได้ให้ความหมายว่า เกณฑ์ คือ ข้อ

กำหนดที่มีขอบเขตบอกระดับพฤติกรรมทั้งหมดในการวัดผลการศึกษาตามวัตถุประสงค์ข้อใด ข้อหนึ่งที่กำหนดไว้ คะแนนที่ได้จากการสอบแต่ละคำมีความหมายบ่งบอกระดับของพฤติกรรม หรือความสามารถของการปฏิบัติของผู้สอบ ส่วนการกำหนดมาตรฐาน หมายถึง การกำหนดระดับ การปฏิบัติเพียงระดับเดียวเป็นจุดตัด (cutting point) ที่แสดงว่าผู้สอบผ่านหรือไม่ผ่าน ได้หรือตก หรือ ควรหรือไม่ควรได้รับประกาศนียบัตร ตามความหมายดังกล่าว ขอบเขตของเกณฑ์แยกระดับ การปฏิบัติงานของผู้สอบออกเป็น 2 กลุ่ม แต่ในทางปฏิบัติ การกำหนดเกณฑ์อาจแยกระดับการ ปฏิบัติของผู้สอบออกเป็นหลายกลุ่ม เช่น แยกออกเป็นเกรด A , B , C , D และ F เป็นต้น

ดิక్షนารีของ Webster (Webster's dictionary, 1984) อธิบายว่า คำว่า เกณฑ์ ในภาษาอังกฤษ คือ criterion - เอกพจน์ , criteria - พหูพจน์ มาจากคำภาษากรีกว่า kriterion = to judge, to decide และให้ความหมายว่า เกณฑ์หมายถึงมาตรฐาน (standards) ซึ่งใช้เป็นหลักในการพิจารณา ตัดสิน (judgement) หรือการตัดสินใจ (decision) ส่วนคำว่ามาตรฐาน (standard) หมายถึงสิ่งที่ กำหนดขึ้นโดยผู้มีอำนาจหน้าที่รับผิดชอบ ไว้เป็นกฎเกณฑ์สำหรับการตรวจสอบคุณภาพ หรือ คุณค่าของสิ่งที่ประเมิน

พัทวิ จริตธรรม (2539) สรุปว่า เกณฑ์ หมายถึงข้อกำหนดที่ใช้เป็นบรรทัดฐานในการตัดสิน คุณภาพหรือพฤติกรรม หรือคุณลักษณะที่ต้องการวัด เกณฑ์ประกอบด้วยมาตรของคะแนนที่เป็น บรรทัดฐานแทนรายการพฤติกรรม หรือคุณลักษณะที่ต้องการวัด มีขอบเขตครอบคลุมระดับ พฤติกรรมหรือคุณลักษณะต่ำสุด ไปจนถึงระดับพฤติกรรมหรือคุณลักษณะสูงสุดตามวัตถุประสงค์ ของการวัด เกณฑ์ที่พัฒนามันต้องมีคะแนนจุดตัดหรือมาตรฐาน (standard) เพื่อแสดงระดับของ พฤติกรรมขั้นต่ำที่ต้องการ

ในที่นี้ ผู้วิจัยมีประเด็นการประเมินสำหรับการวิจัยครั้งนี้เป็นบรรทัดฐาน หรือรายการ พฤติกรรมแล้ว เกณฑ์การประเมินในการวิจัยครั้งนี้ หมายถึง คะแนนที่เป็นมาตรฐานหรือเป็นจุดตัด ที่บ่งบอกระดับคุณภาพขั้นต่ำของตัวเรา

การกำหนดมาตรฐาน (setting standard) หรือคะแนนจุดตัด (cutting point) ทำได้หลายวิธี Glass, G. V. (1978 อ้างถึงใน พัทวิ จริตธรรม, 2539) สรุปว่า วิธีการกำหนดมาตรฐานทำได้ 6 กลุ่มใหญ่ๆ คือ การกำหนดมาตรฐานตามผลการปฏิบัติของกลุ่ม การกำหนดมาตรฐานโดยผู้ เชี่ยวชาญโดยนับถนมาตรฐานลงจาก 100% การกำหนดมาตรฐานวิธีบูทสเตรปปิง (bootstrapping)

ของเกณฑ์ภายนอก การกำหนดมาตรฐานจากสมรรถภาพขั้นต่ำสุด การกำหนดมาตรฐานโดยวิธี การใช้ทฤษฎีการตัดสินใจ และการกำหนดมาตรฐานโดยวิธีวิจัยปฏิบัติการ ในจำนวนวิธีกำหนด มาตรฐานทั้ง 6 วิธีนี้ วิธีแรกเป็นวิธีง่ายและสะดวกสำหรับการประเมิน เพราะใช้ผลการประเมิน โดยอิงค่านิยมฐานของกลุ่มเป็นเกณฑ์ เมื่อนำมาประยุกต์ใช้กับการประเมินตำราอาจใช้เกณฑ์ที่ว่า ผู้ประเมินร้อยละ 50 เห็นว่าคุณภาพของตำราอยู่ในเกณฑ์ดีเป็นมาตรฐานในการตัดสินใจได้

ตอนที่ 4 การพัฒนาตัวบ่งชี้ (Indicators)

เนื่องจากการประเมินคุณภาพตำราสถิติในครั้งนี้ ผู้วิจัยได้กำหนดประเด็นการประเมินไว้ 11 ประเด็น และในแต่ละประเด็นมีประเด็นย่อยรวม 104 ประเด็น เพื่อให้เห็นภาพรวมของผลการ ประเมิน ผลการประเมินจึงน่าจะต้องการพัฒนาตัวบ่งชี้รวมจากประเด็นการประเมินย่อยทั้งหมด ด้วยเหตุนี้ ผู้วิจัยจึงได้ศึกษาวิธีวิทยาการวิจัยเกี่ยวกับการพัฒนาตัวบ่งชี้ และนำเสนอสาระในหัวข้อ นี้

คำว่าตัวบ่งชี้ (indicator) และดัชนี (index) มีความหมายใกล้เคียงกัน Neuman, W. L. (1991 , 146) ให้คำจำกัดความคำว่า ดัชนี คือ ตัวแปรประกอบ (composite variable) ที่เกิดจากการ รวมตัวแปรย่อยหลายตัวแปร ซึ่งให้สารสนเทศเกี่ยวกับสิ่งที่ผู้พัฒนาดัชนีต้องการศึกษา การพัฒนา ดัชนีจึงเป็นกระบวนการลดปริมาณข้อมูล หรือจำนวนตัวแปรให้มีจำนวนน้อยลง ดัชนีจะมีค่า ความเที่ยงและความตรงของการวัดสูงกว่าความเที่ยงและความตรงของตัวแปรที่วัดได้ ค่าของดัชนี อยู่ในรูปข้อมูลเชิงปริมาณที่มีการเปรียบเทียบกับค่าที่ใช้เป็นฐาน และมีค่าต่ำสุดและสูงสุด ตัวแปร ย่อยที่จะนำมาสร้างดัชนี ควรจะมีคุณสมบัติในการวัดคุณลักษณะของสิ่งที่วัดได้อย่างครอบคลุม ครบถ้วน โดยที่ตัวแปรแต่ละตัววัดคุณลักษณะตัวแปรที่ไม่คาบเกี่ยวหรือไม่เกี่ยวข้องกัน (mutually exclusive) นอกจากนี้ตัวแปรที่วัด จะต้องมิใช่ช่วงค่าของตัวแปรครอบคลุมคุณลักษณะทุกประเภทที่ ต้องการวัด (exhaustive) และเมื่อนำตัวแปรย่อยเหล่านี้มาพัฒนาเป็นดัชนีแล้ว ดัชนีควรจะเป็น เอกมิติ (unidimensionality)

Johnstone, J. N. (1981) ให้คำจำกัดความว่า ตัวบ่งชี้ เป็นตัวแปรประกอบ (composite variable) ที่สร้างขึ้นจากการรวมตัวแปรย่อยๆหลายตัวแปร โดยที่ตัวบ่งชี้อาจมีค่าในรูปข้อมูลเชิง ปริมาณหรือเชิงคุณภาพก็ได้ ตามความหมายดังกล่าว ตัวบ่งชี้ จึงมีความหมายกว้างครอบคลุม ความหมายของดัชนี และนักวิชาการหลายท่านใช้คำว่าดัชนี และตัวบ่งชี้ในความหมายเดียวกัน

สำหรับขั้นตอนการพัฒนาดัชนี หรือตัวบ่งชี้ Neuman, W. L. (1991) และ Johnstone, J. N. (1981) สรุปว่า การพัฒนาดัชนีหรือตัวบ่งชี้มีกระบวนการแยกได้เป็น 5 ขั้นตอน คือ

1. การกำหนดจุดมุ่งหมายการพัฒนาตัวบ่งชี้ ตามแนวคิดของ Johnstone, J. N. (1981) ตัวบ่งชี้ที่พัฒนานั้น สามารถนำไปใช้ประโยชน์ในการวางแผน ติดตามกำกับดูแล ตลอดจนประเมินผลจัดการศึกษา Neuman, W. L. (1991) สรุปว่าตัวบ่งชี้ที่พัฒนาขึ้น จะให้ภาพรวมในสิ่งที่ต้องการศึกษา และสามารถนำไปใช้ในการศึกษาความก้าวหน้าหรือพัฒนา หรือการเจริญเติบโตของสิ่งนั้น รวมทั้งการเปรียบเทียบระดับการพัฒนาของสิ่งนั้นๆด้วย

2. การเลือกตัวแปรนำมาสร้างตัวบ่งชี้ ในขั้นตอนนี้ เป็นการพิจารณาตัดสินใจว่าจะเลือกตัวแปรย่อยชนิดใดและจำนวนเท่าไรในการสร้างตัวแปร Johnstone, J. N. (1981) สรุปว่า การตัดสินใจเลือกตัวแปร อาจทำได้ 2 วิธี คือ การเลือกตัวแปรตามแนวคิดทฤษฎีหนึ่ง และการเลือกตัวแปรจากหลักฐานเชิงประจักษ์

3. การสร้างตัวบ่งชี้ ขั้นตอนนี้เป็นการพิจารณาตัดสินใจ กำหนดน้ำหนักความสำคัญ (weighting) แต่ละตัวแปร และพิจารณาว่าจะรวมตัวแปรด้วยวิธีการใด Johnstone, J. N. (1981) และ Neuman, W. L. (1991) เสนอแนะวิธีการสร้างตัวบ่งชี้ไว้หลายแบบ เช่น การหาค่าเฉลี่ย การหาอัตราการเปลี่ยนแปลงการรวมตัวแปรโดยกำหนดน้ำหนักความสำคัญของตัวแปรตามความเห็นผู้เชี่ยวชาญ และการรวมตัวแปรโดยวิธีการวิเคราะห์องค์ประกอบ (factor analysis) ในจำนวนวิธีการพิจารณาตัวบ่งชี้ทั้งหมด Johnstone, J. N. (1981) กล่าวว่า การวิเคราะห์องค์ประกอบโดยใช้ข้อมูลเชิงประจักษ์เป็นวิธีที่ได้รับความนิยมมากที่สุด

การวิเคราะห์องค์ประกอบ เป็นเทคนิคทางสถิติที่มีวัตถุประสงค์ในการลดปริมาณของตัวแปรให้มีจำนวนน้อยลง (variable reduction) โดยอาศัยโครงสร้างและแบบแผนของความสัมพันธ์ที่มีอยู่ในข้อมูลหรือระหว่างตัวแปรเพื่อให้ง่ายต่อความเข้าใจ ซึ่งมีจุดมุ่งหมายสำคัญ 2 ประการ (นงลักษณ์ วิรัชชัย, 2538; อุทุมพร จามรนาน, 2532) จุดมุ่งหมายประการแรก เป็นการสำรวจ และหาองค์ประกอบที่หาความสัมพันธ์ระหว่างตัวแปร เรียกว่าการวิเคราะห์องค์ประกอบเชิงสำรวจ (exploratory factor analysis) วิธีนี้ใช้เมื่อนักวิจัยไม่มีทฤษฎีชัดเจนที่จะใช้เป็นแนวความคิดในการสร้างองค์ประกอบ จุดมุ่งหมายประการที่สอง เป็นการพิสูจน์ สนับสนุน ตรวจสอบสมมุติฐานเกี่ยวกับโครงสร้างของข้อมูล หรือตัวแปรว่ามีองค์ประกอบร่วมกันกี่องค์ประกอบ และองค์ประกอบสัมพันธ์กับตัวประกอบอย่างไร เรียกว่า การวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) วิธีนี้ใช้ในกรณีที่นักวิจัยมีทฤษฎีรองรับ และต้องการตรวจสอบว่าข้อมูลเชิงประจักษ์สอดคล้องเป็นไปตามทฤษฎีเพียงไร การวิเคราะห์องค์ประกอบทั้ง

สองวิธีคือใช้เมทริกซ์สหสัมพันธ์ระหว่างตัวแปรย่อยในการวิเคราะห์ข้อมูลเหมือนกัน

ในการวิเคราะห์องค์ประกอบเชิงสำรวจมีวิธีการสำคัญ 2 ขั้นตอน คือ การสกัดองค์ประกอบ (factor extraction) และการหมุนแกนองค์ประกอบ (factor rotation) การสกัดองค์ประกอบ มีวิธีการทำได้หลายวิธี ซึ่ง Kim, J. O. และ Mueller, C. W. (อ้างถึงในนงลักษณ์ วิรัชชัย, 2538) ได้แยกเป็น 6 กลุ่ม คือ วิธีการวิเคราะห์ส่วนประกอบमुखสำคัญ (principal component analysis) วิธีการหาองค์ประกอบแกนमुखสำคัญ (Principal Axis Factoring) วิธีการกำลังสองน้อยที่สุด (Least Squares Method) วิธีการไลค์ลิตูดสูงสุด (Maximum Likelihood Method) วิวิวิเคราะห์ภาพ (Image Analysis) และวิธีการหาองค์ประกอบแบบแอลฟา (Alpha) ในการสกัดองค์ประกอบมักจะพยายามจะให้ได้ความแปรปรวนมากที่สุดสำหรับองค์ประกอบแต่ละตัว หลังจากที่สกัดองค์ประกอบร่วมของตัวแปรต่างๆ ได้แล้ว ก็จะทราบว่าองค์ประกอบร่วมที่ได้ประกอบด้วยตัวแปรใด โดยดูจากเมทริกซ์น้ำหนักองค์ประกอบซึ่งชี้ให้เห็นถึงความสัมพันธ์ระหว่างองค์ประกอบแต่ละองค์ประกอบกับตัวแปรต่างๆ เหล่านั้น องค์ประกอบที่สกัดได้ก่อนการหมุนแกนในบางครั้ง ก็ยากแก่การอ่านและตีความหมาย ดังนั้นหลังจากการสกัดองค์ประกอบแล้ว จึงต้องมีการหมุนแกน เพื่อให้ได้เมทริกซ์น้ำหนักองค์ประกอบที่มีความหมายมากขึ้น วัตถุประสงค์ที่สำคัญประการหนึ่งของการวิเคราะห์องค์ประกอบ คือ การหาองค์ประกอบที่มีความหมาย องค์ประกอบที่ได้จะมีความหมายชัดเจน เมื่อประกอบด้วยตัวแปรที่มีความสัมพันธ์กัน และมีน้ำหนักองค์ประกอบมากต่อองค์ประกอบนั้น

การหมุนแกนองค์ประกอบ (factor rotation) เป็นวิธีการหนึ่งที่จะทำให้องค์ประกอบมีความหมาย คือ ผลจากการหมุนแกนองค์ประกอบจะทำให้ตัวแปรบางตัวที่แต่เดิมเป็นสมาชิกของหลายองค์ประกอบ กลายเป็นสมาชิกขององค์ประกอบใดองค์ประกอบหนึ่งอย่างเด่นชัดมากขึ้นกว่าเดิม โดยดูได้จากน้ำหนักองค์ประกอบ (factor loading) ของตัวแปรนั้น การหมุนแกนไม่ได้ทำให้ค่าความร่วมกัน (communalities) หรืออัตราส่วนร้อยละของค่าการผันแปรทั้งหมดที่อธิบายโดยองค์ประกอบเปลี่ยนแปลงไป การหมุนแกนเป็นเพียงการเปลี่ยนตำแหน่งของตัวแปรให้สัมพันธ์กับองค์ประกอบในลักษณะที่ชัดเจนขึ้น

การหมุนแกนโดยทั่วไป มี 2 แบบ (นงลักษณ์ วิรัชชัย, 2538 ; อุทุมพร จามรมาน, 2531) คือ การหมุนแกนแบบตั้งฉาก (orthogonal rotation) การหมุนแบบมุมแหลม (oblique rotation) การหมุนแกนแบบตั้งฉาก คือ การหมุนแกนที่องค์ประกอบมีความอิสระต่อกันหรือตั้งฉากซึ่งกันและกัน วิธีการหมุนแกนทำได้ 3 วิธี วิธีแรก คือ วิธีการหมุนแบบแวร์วิแม็กซ์ (varimax rotation)

เป็นวิธีการลดจำนวนตัวแปรในแต่ละองค์ประกอบ ได้องค์ประกอบเฉพาะ (specific factor) โดยทำค่าความแปรปรวนของน้ำหนักองค์ประกอบให้สูงสุดแต่ละองค์ประกอบและพิจารณาเฉพาะตัวแปรที่มีน้ำหนักสูงเท่านั้น วิธีที่สอง คือ วิธีการหมุนแบบควอร์ติแมกซ์ (quartimax rotation) ซึ่งเป็นวิธีการลดจำนวนองค์ประกอบ ได้ผลการวิเคราะห์เป็นองค์ประกอบทั่วไป (general factor) ซึ่งรวมตัวแปรที่สัมพันธ์กันไว้ในองค์ประกอบเดียวกันให้มากที่สุด และวิธีที่สาม คือ วิธีการหมุนแบบอีควอแมกซ์ (equamax rotation) เป็นวิธีการลดจำนวนตัวแปรและองค์ประกอบ เป็นวิธีผสมผสานระหว่างสองวิธีแรก ส่วนการหมุนแบบมุมแหลม เป็นการหมุนแกนที่แกนองค์ประกอบไม่ต้องตั้งฉากกัน หรือแกนทั้งสองทำมุมแหลมต่อกัน แสดงให้เห็นว่าองค์ประกอบมีความสัมพันธ์กัน วิธีการหมุนแกนทำได้ 3 วิธี วิธีแรก คือ วิธีการหมุนแกนแบบควอร์ติมิน (quartimin rotation) ซึ่งเป็นวิธีใช้หลักการเดียวกับวิธีการหมุนแกนแบบควอร์ติแมกซ์ วิธีที่สอง คือ วิธีการหมุนแกนแบบโคแวริมิน (covarimin rotation) ซึ่งเป็นวิธีใช้หลักการเดียวกับวิธีการหมุนแกนแบบเวริแมกซ์ และวิธีที่สาม คือ วิธีการหมุนแกนแบบออบลิมิน (oblimin rotation) ซึ่งเป็นวิธีผสมผสานระหว่างสองวิธีแรก

4. การตรวจสอบความตรงของตัวบ่งชี้ ขั้นตอนนี้เป็นกรนำตัวบ่งชี้ที่พัฒนาไปตรวจสอบว่าตัวบ่งชี้วัดตรงตามคุณลักษณะที่ต้องการวัดหรือไม่ การตรวจสอบความตรงของตัวบ่งชี้ทำได้ทั้งการตรวจสอบความตรงเชิงโครงสร้าง (construct validity) และการตรวจสอบความตรงเชิงเนื้อหา (content validity) รวมทั้งการตรวจสอบความตรงเชิงสภาพ (face validity) สำหรับวิธีการตรวจสอบความตรงเชิงโครงสร้างนี้ นิยมใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน (confirmatory factor analysis) ส่วนการตรวจสอบความตรงเชิงเนื้อหา และความตรงเชิงสภาพมักใช้วิธีการให้ผู้เชี่ยวชาญช่วยการตรวจสอบ

5. การนำเสนอตัวบ่งชี้ Johnstone, J. N. (1981) สรุปว่า เมื่อมีการพัฒนาและตรวจสอบตัวบ่งชี้แล้ว การนำเสนอตัวบ่งชี้ทำได้หลายแบบ เช่น การเสนอในรูปแบบแผนภาพแสดงระดับของตัวบ่งชี้แต่ละตัว หรือเสนอเป็นโปรไฟล์ (profile) การเสนอในรูปแบบการแยกกลุ่มของสิ่งที่ศึกษาตามค่าของตัวบ่งชี้เป็นกลุ่มๆ ด้วยวิธีการวิเคราะห์จำแนกกลุ่ม (cluster analysis)

จากการศึกษาเอกสาร ที่เกี่ยวข้องกับการพัฒนาคณิตหรือตัวบ่งชี้ดังที่กล่าวมาข้างต้น ผู้วิจัยเห็นว่า ในการประเมินคุณภาพโดยการเก็บรวบรวมข้อมูลตามความคิดเห็นของผู้ประเมินในประเด็นใหญ่ๆ 11 ประเด็น และประเด็นย่อยๆ 104 ประเด็น จะทำให้ได้ตัวแปรที่เป็นผลของการประเมินจำนวนมาก ดังนั้นเพื่อให้ได้ภาพที่ชัดเจนเกี่ยวกับคำราสถิติ จึงควรที่จะมีการพัฒนาตัวบ่งชี้ผลการประเมินคุณภาพคำราสถิติในแต่ละประเด็นด้วย สำหรับวิธีการพัฒนาตัวบ่งชี้ในการ

วิจัยครั้งนี้ ผู้วิจัยพิจารณาเห็นว่า ประเด็นการประเมินทั้ง 11 ประเด็นนั้นมีความเกี่ยวข้องกับสัมพันธ์กัน จึงมุ่งพัฒนาดัชนีหรือตัวบ่งชี้จากประเด็นการประเมินทั้ง 11 ประเด็น โดยใช้การวิเคราะห์องค์ประกอบแบบสำรวจ

ตอนที่ 5 แนวทางการประเมินตำราสถิติ

จากการศึกษารายงานเอกสารและงานวิจัยที่เกี่ยวข้องกับการวิจัยครั้งนี้ ผู้วิจัยสรุปเป็นแนวทางการวิเคราะห์เชิงประเมินตำราสถิติที่ดำเนินการวิจัย ดังแผนภูมิที่ 1

แนวทางการวิเคราะห์เชิงประเมินตำราสถิติ มีขั้นตอนดำเนินการดังนี้

ขั้นตอนที่ 1 กำหนดวัตถุประสงค์การวิเคราะห์เชิงประเมินตำราสถิติ ครั้งนี้ ผู้วิจัยมีวัตถุประสงค์เพื่อประเมินคุณภาพตำราสถิติที่คัดสรรจำนวน 3 เล่ม คือ ตำราสถิติเล่ม ก ตำราสถิติเล่ม ข และตำราสถิติเล่ม ค และเพื่อพัฒนาตัวบ่งชี้รวมบ่งบอกคุณภาพของตำราสถิติคัดสรรทั้ง 3 เล่ม

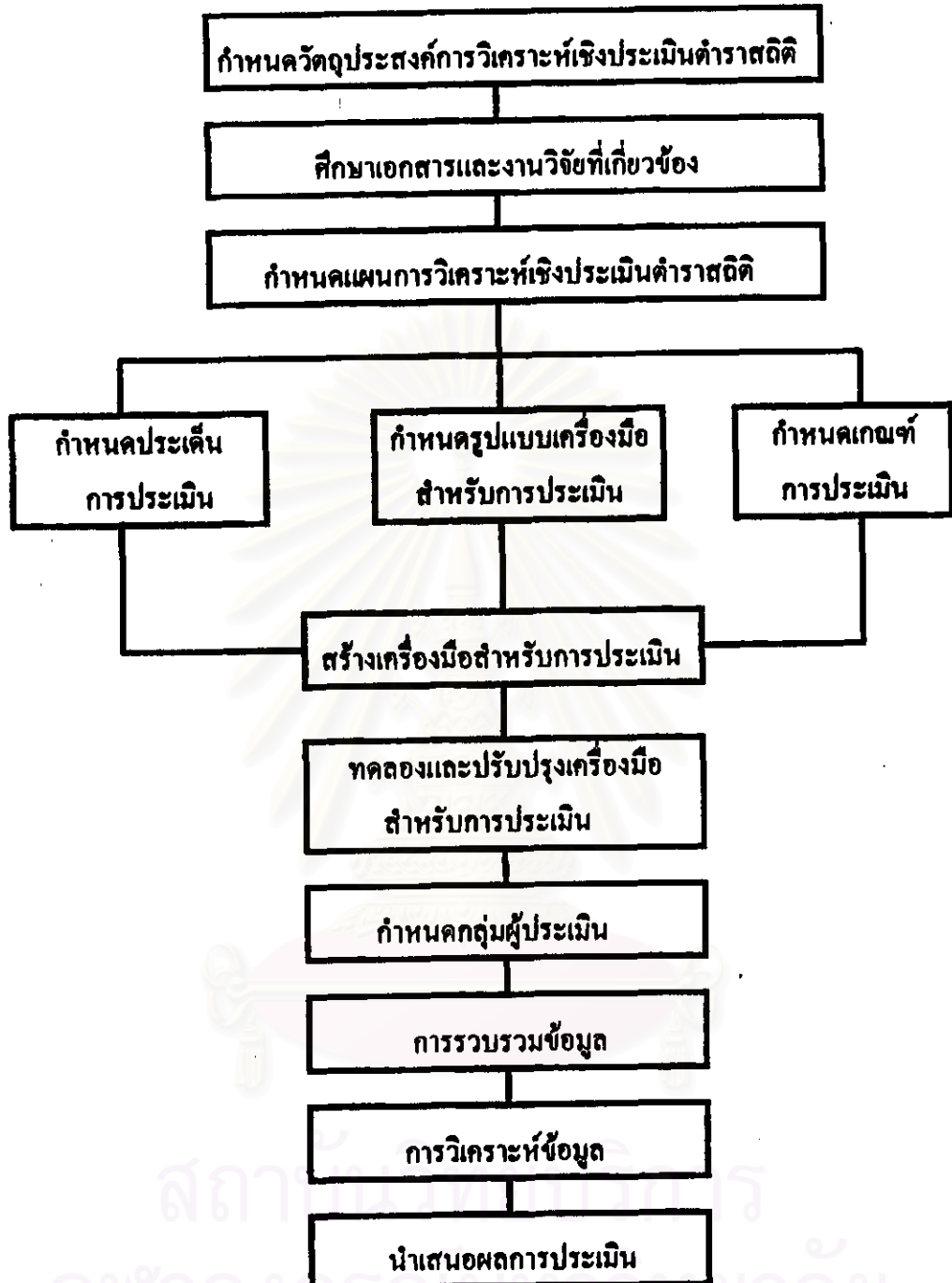
ขั้นตอนที่ 2 ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ในขั้นตอนนี้ ผู้วิจัยศึกษาจากผลงานวิจัยเชิงประเมินตำราสถิติจำนวน 7 เรื่อง ได้แก่ งานวิจัยของ Harwell, M. R. และคณะ (1996) ; Huberty, C. J. และ Barton, R. M. (1990) ; Cobb, G. W. (1987) ; และ Berry, K. และ Lee, D. (1982) ผลงานของ Hollabaugh, M. (1989) ; และ Jacobson, M. D. (1965)

ขั้นตอนที่ 3 กำหนดแผนการวิเคราะห์เชิงประเมินตำราสถิติ จากขั้นตอนที่ 2 ผู้วิจัยวางแนวทางการประเมินและกำหนดระยะเวลาในการดำเนินการวิจัยของแต่ละขั้นตอน

ขั้นตอนที่ 4 การสร้างเครื่องมือสำหรับการประเมิน จากขั้นตอนที่ 2 ผู้วิจัยได้ข้อสรุปเป็นวิธีการประเมินตำราในครั้งนี้ คือ กำหนดประเด็นการประเมิน กำหนดรูปแบบเครื่องมือสำหรับการประเมิน และกำหนดเกณฑ์การประเมิน ดังนี้

ก. กำหนดประเด็นการประเมิน จากการศึกษาเอกสาร งานวิจัย ผู้ประเมินได้แนวคิดประเด็นการประเมินตำราสถิติในครั้งนี้ รวม 11 ด้าน คือ ลักษณะรูปเล่ม รูปแบบการเขียน ความน่าอ่าน วิธีการนำเสนอเนื้อหา ความครอบคลุมของเนื้อหา ความถูกต้องของเนื้อหา ตารางของตำราที่ทำให้เข้าใจดีขึ้น ลักษณะของแบบฝึกหัด ส่วนเสริมของตำรา การให้คำแนะนำในการสอน และ ด้านลักษณะผู้แต่งตำรา

ข. กำหนดรูปแบบเครื่องมือสำหรับการประเมิน ผู้วิจัยได้ใช้แบบประเมินที่สร้างขึ้นมี 2 ชุด คือแบบประเมินสำหรับนิสิต และแบบประเมินสำหรับผู้สอน มีลักษณะเป็นมาตราประเมินค่าแบบลิเคอร์ท (Likert) ให้คะแนน 5 ระดับ



แผนภูมิที่ 1 ขั้นตอนในการดำเนินการวิเคราะห์เชิงประเมิน

ก. กำหนดเกณฑ์การประเมิน ผู้วิจัยใช้เกณฑ์การประเมินอย่างง่าย จากมาตรฐานตามผลการปฏิบัติของกลุ่ม โดยอิงคำมัธยฐานของกลุ่มเป็นเกณฑ์ ในการวิจัยครั้งนี้ ผู้วิจัยกำหนดเกณฑ์ผู้ประเมินร้อยละ 50 มีความเห็นว่าคุณภาพของประเด็นการประเมินอยู่ในระดับดี ถือว่าคุณภาพของตำราเป็นที่ยอมรับได้

ขั้นตอนที่ 5 ทดลองและปรับปรุงเครื่องมือสำหรับการประเมิน ผู้วิจัยนำเครื่องมือที่สร้างขึ้นไปทดลองใช้กับกลุ่มตัวอย่างทดลอง เพื่อตรวจสอบและปรับปรุงแก้ไขเครื่องมือ

ขั้นตอนที่ 6 กำหนดกลุ่มผู้ประเมิน เป็นการพิจารณากำหนดกลุ่มตัวอย่างในการวิจัย ในที่นี้คืออาจารย์ผู้สอนวิชาสถิติประยุกต์ทางพฤติกรรมศาสตร์ รหัสวิชา 2702601 และ นิติศัลกัศูตรปริญญาคุณวุฒิจิตและมหาบัณฑิต คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ขั้นตอนที่ 7 การรวบรวมข้อมูล ผู้วิจัยกำหนดระยะเวลาของการเก็บรวบรวมข้อมูลตั้งแต่ปลายเดือนมีนาคม ถึง เดือนกันยายน 2540

ขั้นตอนที่ 8 การวิเคราะห์ข้อมูล ผู้วิจัยใช้สถิติบรรยาย เสนอเป็นคำร้อยละของข้อมูลที่เป็นภูมิหลัง และการวิเคราะห์ข้อมูลที่เป็นผลการประเมินจากแบบประเมินคำร่าเป็นคำสถิติพื้นฐาน

ขั้นตอนที่ 9 การนำเสนอผลการประเมิน เป็นการสรุปผลการวิจัย ที่ค้นพบในการวิจัยครั้งนี้ โดยอภิปรายผลที่ได้ ข้อเสนอแนะในการนำผลเอาไปใช้ และข้อเสนอแนะในการทำวิจัยครั้งต่อไป



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย