การใช้สถิติในงานวิจัยทางภาษา

Statistical Application in Linguistic Studies

ผจงจิต อินทสุวรรณ

Pachongchit Intasuwan

บทคัดย่อ

งานวิจัยเชิงปริมาณทางภาษาต้องมีการวิเคราะห์ข้อมูลด้วยสถิติเพื่อให้ได้ข้อสรุปที่เป็นระบบและมีความหมายชัดเจน สถิติที่ใช้มีทั้งสถิติบรรยายและสถิติอนุมาน สถิติบรรยายเป็นการอธิบายลักษณะของตัวแปรที่ศึกษาในกลุ่มประชากร ส่วนสถิติอนุมานเป็นการสรุปเกี่ยวกับตัวแปรที่ศึกษาในกลุ่มประชากรโดยอาศัยข้อมูลที่รวบรวมจากกลุ่มตัวอย่างและผ่านกระบวนการทดสอบสมมุติฐาน สถิติอนุมานที่กล่าวถึงในที่นี้เป็นสถิติที่พบบ่อย ๆ ในงานวิจัย อันได้แก่ สถิติที (t-test) การวิเคราะห์ความแปรปรวน (ANOVA) การวิเคราะห์สหสัมพันธ์ (correlation analysis) และสถิติไค-สแควร์ (chi-square test) สถิติสามตัวแรกอยู่ในกลุ่มของสถิติที่ต้องมีเงื่อนไขเกี่ยวกับการแจกแจงของประชากร (parametric statistics) ส่วนสถิติตัวสุดท้ายอยู่ในกลุ่ม non-parametric statistics ซึ่งไม่ต้องมีเงื่อนไขเกี่ยวกับการแจกแจงของประชากร โดยทั่วไปสถิติกลุ่มแรกมีอำนาจ (power) สูงกว่ากลุ่มหลัง การเลือกใช้สถิติเป็นไปตามลักษณะของข้อมูลและคำถามการวิจัย

Abstract

Quantitative data in language research need a statistical analysis to give a clear, systematic, and meaningful conclusion. The statistics used in the research are both descriptive and inferential. While the descriptive statistics explain the population characteristics, the inferential ones provide conclusions for populations, using sample data, through a hypothesis testing. Commonly used statistics described in this paper are t-test, analysis of variance (ANOVA), correlation analysis, and chi-square test. The first three tests are parametric-statistics which need to assume the shape of the population distributions. The last test is non-parametric which needs no assumption of the population distribution. In general, the former type of statistics is more powerful than the

last one. However, the choice of statistics to be used in each case depends on the measurement scale and the research questions.

ความนำ

แม้ว่างานวิจัยทางภาษาจะใช้วิธีการบรรยายกันเป็นจำนวนมาก เช่นงานวิจัยเรื่อง "การสังเคราะห์งานวิจัยทางภาษาศาสตร์เชิงจิตวิทยา" (วรรณา ฐิติโรจน์ไพบูลย์, 2536) "ลักษณะภาษาโฆษณาสำหรับผู้หญิงในสื่อนิตยสารตั้งแต่ปี พ.ศ. 2475 ถึง พ.ศ. 2543" (วัสสิกา รุมาคม, 2545)"วิเคราะห์ท่วงทำนองการประพันธ์ในสุภาษิตพระร่วง" (อัญชลี วรัญญา บริรักษ์, 2546) "การศึกษารูปแบบภาษานิยมในบทสนทนาภาษาไทย" (วิศนี เครือวนิชกรกุล, 2547) และ "Understanding changes in elementary Mandarin students' L1 and L2 writing" (McCarthey, Guo, & Cummins, 2005) เป็นต้น แต่ก็มีงานวิจัยทางภาษาอีกเป็นจำนวนมากได้ใช้วิธีวิเคราะห์ข้อมูลเชิงปริมาณเพื่อตอบคำถามการวิจัย ซึ่งจะช่วยให้สามารถสรุปภาพรวมของประเด็นที่สนใจศึกษาได้อย่างเป็นระบบที่มีความหมายชัดเจนยิ่งขึ้น

งานวิจัยที่รวบรวมข้อมูลเชิงปริมาณและใช้สถิติวิเคราะห์นั้น มีตั้งแต่การใช้สถิติบรรยาย (descriptive statistics) และสถิติอนุมาน (inferential statistics)

การใช้สถิติบรรยายเป็นการนำเสนอข้อมูลเชิงปริมาณที่มีจำนวนมากให้อยู่ในรูปที่มีความหมายซึ่งคนทั่วไปอ่านเข้าใจง่าย เช่น การรวบรวมคำคุณศัพท์ที่ใช้ในการเขียนจากบทความจำนวนมาก ผู้วิจัยเสนอผลในรูปของการแจกแจงความถี่และร้อยละ เช่น Banpho (2004: 73, 82) พบว่า ในประโยคเริ่มเรื่องมีการใช้ present simple tense บ่อยที่สุด คิดเป็น 70.06% ตามด้วย present perfect tense คิดเป็น 17.37% เป็นต้น หรือรายงานในรูปของค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (standard deviation) เช่น Banpho (2004: 81-82) พบว่า ความยาวของประโยคเริ่มเรื่องในวารสารทางการแพทย์ 2 ฉบับ มีค่าเฉลี่ยเป็น 21.89 คำ และมีส่วนเบี่ยงเบนมาตรฐานเป็น 9.89 เป็นต้น

การใช้สถิติอนุมานเป็นการใช้ผลที่พบจากกลุ่มตัวอย่าง (sample) เพื่อสรุปเป็นลักษณะของประชากร (population) กลุ่มตัวอย่างเป็นตัวแทนส่วนหนึ่งของประชากร ในการวิจัยส่วนใหญ่นักวิจัยไม่สามารถเก็บรวบรวมข้อมูลจากประชากรได้ เช่น ประชากรเป็นนักเรียนชั้นมัธยมศึกษาปีที่ 2 ในเขตกรุงเทพมหานคร ผู้วิจัยสนใจคะแนนภาษาอังกฤษของนักเรียนเหล่านี้ เนื่องจากนักเรียนมีจำนวนมาก ดังนั้นจึงเลือกตัวอย่างมากลุ่มหนึ่ง แล้วรวบรวมข้อมูลที่สนใจจากกลุ่มตัวอย่างนี้ (ผู้วิจัยต้องคำนึงถึงหลักการที่ถูกต้องในการได้กลุ่มตัวอย่างที่เป็นตัวแทนของประชากร) ซึ่งประหยัดเวลาและการลงทุน กลุ่มตัวอย่างอาจประกอบไปด้วยนักเรียนจำนวน 200 คน ผู้วิจัยทดสอบความรู้ภาษาอังกฤษจากนักเรียนในกลุ่มตัวอย่างนี้ คำนวณค่าเฉลี่ย ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐาน (SD) และใช้หลักการทางสถิติในการสรุปว่าค่าสถิติที่คำนวณได้จากกลุ่มตัวอย่างเป็นค่าพารามิเตอร์ของประชากร

ในบทความนี้จะกล่าวถึงความหมายและตัวอย่างของการใช้สถิติอนุมานในงานวิจัย ๑
ทางภาษาศาสตร์ ที่มีการตั้งสมมุติฐานการวิจัย โดยเลือกกล่าวถึงสถิติที่ใช้กันบ่อย ๆ คือ การทดสอบที (t-test) การวิเคราะห์ความแปรปรวน (analysis of variance - ANOVA) สหสัมพันธ์ (correlation - r) และการทดสอบไค-สแควร์ (chi-square - χ^2) โดยไม่มีรายละเอียดเกี่ยวกับสูตรในการคำนวณ เนื่องจากผู้วิจัยสามารถใช้บริการที่สะดวกจากโปรแกรมคอมพิวเตอร์ได้ เช่น โปรแกรม SPSS ซึ่งมีหนังสือเขียนแนะนำวิธีใช้ไว้อย่างดี เช่น Field (2000)

ก่อนจะกล่าวถึงการใช้สถิติในการทดสอบสมมุติฐาน ควรกล่าวโดยย่อถึงความหมายของสมมุติฐานก่อน ส่วนรายละเอียดสามารถอ่านในหนังสือสถิติทั่ว ๆ ไป เช่น ผจงจิต อินทสุวรรณ (2528) และ Wilcox (1996) เป็นต้น

ในการวิจัยผู้วิจัยอาจตั้งสมมุติฐานการวิจัย (research hypothesis) เช่น การเปรียบเทียบความรู้ภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 2 ระหว่างกลุ่มนักเรียนชายกับกลุ่มนักเรียนหญิง อาจตั้งสมมุติฐานการวิจัยว่า "มีความแตกต่างระหว่างกลุ่มทั้งสอง" นี่เป็นการทดสอบสองทาง (two-sided test) หรือในการทดลองสอนคำศัพท์ให้กับกลุ่มทดลองและเปรียบเทียบกับกลุ่มควบคุม (ซึ่งไม่ได้เน้นการสอนคำศัพท์) ผู้วิจัยอาจตั้งสมมุติฐานการวิจัยว่า "กลุ่มทดลองมีคะแนนคำศัพท์สูงกว่ากลุ่มควบคุม" นี่เป็นการทดสอบทางเดียว (one-sided test) เมื่อจะทดสอบสมมุติฐานผู้วิจัยจะต้องเปลี่ยนสมมุติฐานการวิจัยให้เป็นสมมุติฐานทางสถิติ ซึ่งประกอบด้วย สมมุติฐานเป็นกลาง (null hypothesis - H_0) และสมมุติฐานทางเลือก (alternative hypothesis - H_1) สำหรับสมมุติฐานทางเลือกนั้นเทียบได้กับสมมุติฐานการวิจัย ส่วนสมมุติฐานเป็นกลาง (H_0) นั้นหมายถึงการไม่มีความแตกต่าง เช่น H_0 สำหรับตัวอย่างแรกจะเป็นข้อความ "ไม่มีความแตกต่างระหว่างกลุ่มชายหญิง" และตัวอย่างที่สองจะเป็น "ไม่มีความแตกต่างระหว่างกลุ่มทดลองและกลุ่มควบคุม" หรือ "กลุ่มทดลองมีคะแนนเท่ากับหรือต่ำกว่ากลุ่มควบคุม"

ในการใช้สถิติทดสอบสมมุติฐาน ผู้วิจัยสมมติว่า H_0 เป็นจริง และทดสอบ H_0 ถ้ามีหลักฐานอย่างใดอย่างหนึ่งที่แสดงว่า H_0 ไม่เป็นจริงแล้ว ผู้วิจัยจะปฏิเสธ H_0 หรือกล่าวว่า H_0 ไม่เป็นจริง หรือกล่าวว่า H_1 เป็นจริงด้วยความคลาดเคลื่อนระดับหนึ่ง (type I error หรือ α) (เช่น 5%) หรือกล่าวได้ว่าความแตกต่างมีนัยสำคัญ (significant difference) แต่ถ้าผู้วิจัยไม่มีหลักฐานปฏิเสธ H_0 ก็จะสรุปได้เพียงว่ายังไม่มีหลักฐานปฏิเสธ H_0 แต่ไม่ได้หมายความว่า H_1 เป็นจริง เพียงแต่ข้อมูลที่มีไม่สามารถสนับสนุนสมมุติฐานการวิจัย ดังนั้นผลการทดสอบสมมุติฐานที่เป็นประโยชน์ในการลงข้อสรุป คือ การปฏิเสธ H_0

ในการทดสอบสมมุติฐานนั้น ผู้วิจัยควรตระหนักว่าไม่สามารถกระทำด้วยความเชื่อมั่น 100% ความคลาดเคลื่อนชนิดที่ควรระวังอย่างมาก คือ ความคลาดเคลื่อนชนิดที่ 1 (Type

I error ซึ่งใช้สัญลักษณ์ α) เป็นค่าความน่าจะเป็นในการปฏิเสธ Ho ในขณะที่ Ho เป็นจริง ผู้วิจัยจะพยายามใช้ α ระดับต่ำมาก เช่น .05 (หรือ 5%) หรือ .01 (หรือ 1%) เนื่องจากเป็นความคลาดเคลื่อนที่ร้ายแรง

ความหมายของสถิติและตัวอย่างที่ใช้

1. การทดสอบที (t-test)

สถิติที เป็นสถิติที่ใช้ทดสอบความแตกต่างระหว่างกลุ่ม 2 กลุ่ม ซึ่งอาจเป็น 2 กลุ่มที่เป็นอิสระต่อกันหรือไม่ก็ได้ กลุ่ม 2 กลุ่มที่เป็นอิสระต่อกัน แสดงว่าแต่ละหน่วยในกลุ่มตัวอย่างสุ่มมาจากประชากรต่างกัน เช่น กลุ่มชาย-หญิง กลุ่มอาชีพรับราชการ-ไม่รับราชการ เป็นต้น ส่วนกลุ่มที่ไม่เป็นอิสระต่อกันนั้นแต่ละหน่วยในกลุ่มตัวอย่างทั้งสองมีความเกี่ยวข้องกัน เช่น กลุ่มคู่แฝด กลุ่มผู้ปกครอง-เด็กในปกครอง คะแนนที่วัดกลุ่มเดียวกันซ้ำ 2 ครั้ง เป็นต้น ผู้วิจัยทดสอบสมมุติฐานเพื่อเป็นการตอบคำถามว่ากลุ่มทั้งสองมาจากประชากรเดียวกันหรือไม่ กลุ่มทั้งสองอาจปรากฏว่าแตกต่างกันเนื่องจากความคลาดเคลื่อน หรือมีความแตกต่างจริง (อย่างมีนัยสำคัญ) ที่เราจะกล่าวได้ว่า กลุ่มทั้งสองมาจากประชากร 2 กลุ่มที่แตกต่างกัน

1.1 การทดสอบทีสำหรับกลุ่มตัวอย่างที่เป็นอิสระต่อกัน (t-test for independent groups)

ในกรณีนี้การทดสอบทีมีเงื่อนไขหรือข้อตกลงเบื้องต้น (assumption) 3 ประการ (Wilcox, 1993, p. 130) คือ

1. คะแนนแต่ละตัว สุ่มมาจากกลุ่มที่เป็นอิสระต่อกัน (random sampling from independent groups)
2. ประชากรทั้ง 2 กลุ่มมีการกระจายเป็นปกติ (normal distribution)
3. ประชากรมีความแปรปรวนเท่ากัน (homogeneity of variance)

ตัวอย่างงานวิจัยที่ใช้สถิติทีสำหรับกลุ่มตัวอย่างที่เป็นอิสระต่อกัน

1. พิณทิพย์ ทวยเจริญ (2543, หน้า 9-31) ได้ทดลองสอนภาษาอังกฤษให้กับเด็กวัย 8-11 ปี เพื่อดูผลการเรียนของเด็กใน 3 มิติ คือ พูด-อ่าน ออกเสียง-ฟัง และเขียน ระหว่างกลุ่มทดลองกับกลุ่มควบคุม ในตาราง 1 แสดงผลการเปรียบเทียบเฉพาะด้านการฟัง(พิณทิพย์ ทวยเจริญ. 2543, หน้า 21)

กลุ่ม	N	$\overline{X}$	SD	t
ทดลอง	37	98.65	8.22	54.12**
ควบคุม	42	1.19	7.72	

** มีนัยสำคัญที่ระดับ .01

ในกรณีเช่นนี้ผู้วิจัยได้ทำการสอนภาษาอังกฤษให้กับเด็กนักเรียน จำนวน 37 คน ซึ่งเป็นกลุ่มทดลอง หลังจากสอนแล้วได้ทดสอบความสามารถในการฟังทั้งนักเรียนในกลุ่มทดลองและกลุ่มควบคุม (นักเรียนที่เรียนตามปกติของโรงเรียน) คะแนนสอบในแต่ละกลุ่มนำมาคำนวณค่าเฉลี่ย ($\overline{X}$) และส่วนเบี่ยงเบนมาตรฐาน (SD) (ค่า $\overline{X}$ ในสองกลุ่มนี้แตกต่างกันอย่างมาก ส่วนค่า SD ใกล้เคียงกันมาก) แล้วนำมาคำนวณค่าสถิติที ซึ่งพบว่ามีนัยสำคัญสูงมาก แสดงว่าคะแนนในกลุ่มทดลองสูงกว่ากลุ่มควบคุมอย่างเด่นชัด ใช้เป็นหลักฐานที่แสดงถึงผลการสอนได้อย่างดี (ในปัจจุบันผู้วิจัยสามารถ ป้อนข้อมูลดิบลงในคอมพิวเตอร์ แล้วใช้โปรแกรมสำเร็จรูปซึ่งมีหลากหลายวิเคราะห์ผลได้โดยสะดวก)

สมมุติฐานทางการวิจัยน่าจะเป็นดังนี้

คะแนนการฟังภาษาอังกฤษในกลุ่มทดลองน่าจะสูงกว่าในกลุ่มควบคุม

ส่วนสมมุติฐานทางสถิติน่าจะเป็นดังนี้ (การทดสอบทางเดียว)

H_0 : คะแนนเฉลี่ยการฟังในกลุ่มทดลองและกลุ่มควบคุมเท่ากัน
หรือ คะแนนกลุ่มทดลองเท่ากับหรือต่ำกว่ากลุ่มควบคุม

H_1 : คะแนนเฉลี่ยการฟังในกลุ่มทดลองสูงกว่ากลุ่มควบคุม

การตัดสินใจปฏิเสธหรือไม่ปฏิเสธ Ho ทำได้ 2 ทาง ทางแรกอ่านค่า p (p-value) จาก computer printout ถ้า p มีค่าน้อยกว่าระดับ α ที่ตั้งไว้ เช่น .05 เราสรุปว่าปฏิเสธ Ho ด้วยระดับนัยสำคัญ .05 ทางที่สองใช้การเปรียบเทียบค่า t ที่คำนวณได้กับค่าทีที่อ่านจากตาราง (Students' t table) ด้วยระดับ α ที่กำหนด และ degrees of freedom (df) ของสถิติ ถ้า t ที่คำนวณได้มีค่าสูงกว่าค่า t ที่อ่านจากตาราง จึงสรุปว่าปฏิเสธ Ho มิฉะนั้นจะไม่ปฏิเสธ

2. งานวิจัยของ Banpho (2004, p.81-82) ได้สำรวจความยาวเฉลี่ยของประโยคเริ่มแรก จากบทความในวารสารทางการแพทย์ 2 ฉบับ (72 และ 95 บทความจากฉบับที่ 1 และฉบับที่ 2 ตามลำดับ) พบว่า ค่าเฉลี่ยเป็น 21.79 และ 21.96 สำหรับฉบับที่ 1 และ 2 ตามลำดับ ค่าของส่วนเบี่ยงเบนมาตรฐานเป็น 10.37 และ 9.52 ตามลำดับ ผู้วิจัยไม่ได้ระบุสมมุติฐานไว้ในบทความ แต่น่าจะเป็นดังนี้ (การทดสอบสองทาง)

H_0 : ค่าเฉลี่ยของความยาวของประโยคใน 2 ฉบับเท่ากัน

H_1 : ค่าเฉลี่ยไม่เท่ากัน

สถิติที่ผู้วิจัยใช้คือ t-test สำหรับ 2 กลุ่มที่เป็นอิสระต่อกัน

ผลการคำนวณได้ค่า t = .107 ซึ่งเป็นค่าที่ไม่มีนัยสำคัญที่ระดับ α = .05 เนื่องจากอ่านค่า t จากตาราง ($_{.975}t_{165}$) ได้ 1.96 แสดงว่าข้อมูลที่รวบรวมได้ ไม่ทำให้ปฏิเสธ Ho นั่นคือ ไม่ยืนยันว่าความยาวของประโยคในวารสาร 2 ฉบับดังกล่าวมีความแตกต่างกัน

1.2 การทดสอบทีสำหรับกลุ่มตัวอย่างที่ไม่เป็นอิสระต่อกัน (t-test for dependent groups)

งานวิจัยเชิงทดลองจำนวนมากมักจะสนใจว่า หลังจากได้จัดกระทำบางอย่างกับกลุ่มตัวอย่างแล้ว จะมีการเปลี่ยนแปลงในตัวแปรที่ผู้วิจัยสนใจหรือไม่ ในกรณีนี้คะแนนที่วัดก่อนและหลังการทดลอง เป็นตัวอย่างหนึ่งของกรณีที่มีคะแนน 2 ชุดที่ไม่เป็นอิสระต่อกัน (การวัดตัวแปรซ้ำกลุ่มตัวอย่างเดิม)

สมมุติฐานที่ทดสอบในกรณีที่กลุ่มตัวอย่างไม่เป็นอิสระกัน คือ

H_0 : ไม่มีความแตกต่างกันระหว่างกลุ่ม (เช่นวัดก่อนวัดหลังไม่แตกต่างกัน)

H_1 : มีความแตกต่างกันระหว่างกลุ่ม (สมมุติฐานไม่มีทิศทาง) หรือ

H_1 : คะแนนที่วัดหลังสูงกว่าที่วัดก่อน (สมมุติฐานมีทิศทาง)

ตัวอย่างงานวิจัยที่ใช้การทดสอบที ชนิดกลุ่มตัวอย่างไม่เป็นอิสระกัน

งานวิจัยของพิณทิพย์ ทวยเจริญ (2543) ซึ่งทดลองสอนภาษาอังกฤษให้กับเด็กในช่วงอายุ 8-11 ปี และเมื่อสอนแล้วต้องการทราบว่า การฟัง (ตัวแปรตาม) ของเด็กหลังสอนสูงกว่าก่อนสอนหรือไม่ (มีทิศทาง) ผลการคำนวณข้อมูลปรากฏดังนี้ (พิณทิพย์ ทวยเจริญ, 2543, หน้า 22)

การวัด	n	$\bar{X}$	SD	t
ก่อน	37	98.65	8.22	73.00**
หลัง	37	0.0	0	

** ค่าสถิติมีนัยสำคัญที่ระดับ .01

เนื่องจากค่า t ที่คำนวณได้นี้สูงกว่าค่า t ที่อ่านจากตาราง ($_{.99}t_{36}$ = 2.437) ดังนั้น ผู้วิจัยสรุปว่า เด็กมีความสามารถในการฟังหลังสอนสูงกว่าก่อนสอนอย่างเด่นชัด

การวิเคราะห์ความแปรปรวน (Analysis of variance – ANOVA)

ANOVA เป็นวิธีทดสอบสมมุติฐานเกี่ยวกับค่าเฉลี่ยเช่นเดียวกับสถิติที แต่เป็นการทดสอบความแตกต่างระหว่างกลุ่ม 2 กลุ่มหรือมากกว่าก็ได้ เช่น แทนที่จะเปรียบเทียบความแตกต่างของคะแนนความรู้ภาษาอังกฤษระหว่างนักเรียนชายกับนักเรียนหญิง (2 กลุ่ม) ซึ่งสามารถทดสอบด้วยสถิติทีหรือ ANOVA ก็ได้ แต่ถ้าเป็นความแตกต่างระหว่าง 3 กลุ่มขึ้นไป เช่น นักเรียนที่บิดามีอาชีพต่างกัน 4 กลุ่ม ในกรณีนี้ต้องใช้ ANOVA ในการเปรียบเทียบ จึงสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้อย่างเหมาะสม

การใช้ ANOVA มีเงื่อนไข 3 ประการเช่นเดียวกับ t-test นั่นคือ ข้อมูลเป็นคะแนนชนิดสุ่ม การแจกแจงเป็นปกติ และความแปรปรวนเท่ากัน

สมมติฐานทางสถิติสำหรับ ANOVA คือ

H_0 : คะแนนเฉลี่ยในแต่ละกลุ่มเท่ากัน (นั่นคือไม่มีความแตกต่างระหว่างกลุ่ม)

H_1 : Ho ไม่จริง (มีอย่างน้อย 1 กลุ่มแตกต่างจากกลุ่มอื่น ๆ)

สถิติที่ใช้ทดสอบสมมุติฐาน คือ การทดสอบค่าเอฟ (F-test) ซึ่งเป็นการใช้อัตราส่วนของความแปรปรวนระหว่างกลุ่มกับความแปรปรวนภายในกลุ่ม กฎการตัดสินใจในการปฏิเสธ Ho (หรือไม่ปฏิเสธ Ho) เป็นเช่นเดียวกับการทดสอบสถิติที

ในที่นี้จะกล่าวเฉพาะ ANOVA ที่มีมิติเดียว หรือตัวแปรอิสระเดียว (one-way ANOVA) ส่วนกรณีที่มีหลายตัวแปร ผู้อ่านสามารถดูได้จากหนังสือสถิติทั่วไปในหัวข้อ Factorial ANOVA

ตัวอย่างการใช้ ANOVA

1. งานวิจัยของพิณทิพย์ ทวยเจริญ (2543) มีวัตถุประสงค์อีกประการหนึ่งคือ เปรียบเทียบพัฒนาการทางภาษาของเด็ก 3 วัย คือ 8;0 - 9;0 ปี 9;1 - 10;0 ปี และ 10;1 - 11 ปี โดยใช้คะแนนความแตกต่างระหว่างก่อนเรียนและหลังเรียนเป็นตัวชี้พัฒนาการ ในกรณีนี้เป็นการเปรียบเทียบ 3 กลุ่มจึงใช้ ANOVA ผลเฉพาะด้านการฟังเป็นดังนี้ (พิณทิพย์ ทวยเจริญ. 2543: 23-24)

Source of variation	df	SS	MS	F
Between group	2	64.01	32.01	.46
Within group	34	2368.42	69.66	
Total	36	2432.43		

อัตราส่วน F = .46 เป็นค่าที่ไม่มีนัยสำคัญ ผลนี้จึงแสดงว่าการสอนภาษาอังกฤษให้กับเด็กวัยใดใน 3 วัยนี้ให้ผลไม่ต่างกัน (การตัดสินใจปฏิเสธ Ho หรือไม่ มี 2 ทาง เช่นเดียวกันกับสถิติที)

สมมุติฐานการวิจัยน่าจะเป็นว่า "พัฒนาการในการฟังของเด็ก 3 วัย น่าจะแตกต่างกัน"

ส่วนสมมุติฐานทางสถิตินำจะเป็นดังนี้

H_0 : พัฒนาการในการฟังของเด็กทั้ง 3 กลุ่มไม่แตกต่างกัน

H_1 : Ho ไม่ถูกต้อง

หรือ พัฒนาการในการฟังของเด็กกลุ่มใดกลุ่มหนึ่งแตกต่างจากกลุ่มอื่น ๆ

2. งานวิจัยของเกศแก้ว เพ็งพริ้ง (2547) เปรียบเทียบความสามารถในการอ่านและเขียนคำถูกอักขรวิธี ในที่นี้จะยกตัวอย่างเฉพาะการอ่าน จำแนกตามกลุ่มแผนการเรียน 3

แผน คือ 1) คณิต-วิทย์ (108 คน) 2) คณิต-อังกฤษ และอังกฤษ-คณิต ก (57 คน) และ 3) ภาษาต่างประเทศ (40 คน) ผลการวิเคราะห์ข้อมูลด้วย ANOVA เป็นดังนี้ (เกศแก้ว เพ็งพริ้ง. 2547: 38)

กลุ่ม	n	$\bar{X}$	SD	F	p-value
1	108	21.01	3.92	15.172	.000
2	57	17.40	4.44		
3	40	19.20	3.76		

พิจารณาจากค่า p-value ที่ปรากฏ .000 (ผู้วิจัยไม่ได้รายงานค่า SS และ MS ไว้) แสดงว่า การทดสอบมีนัยสำคัญสูงมาก เป็นการยืนยันว่าค่าเฉลี่ย 3 กลุ่มนี้แตกต่างกันอย่างเด่นชัด ผู้วิจัยจึงเปรียบเทียบรายคู่ต่อไปด้วยวิธีการของ Scheffe' และได้พบว่า กลุ่มที่ 1 และ 2 มีคะแนนเฉลี่ยแตกต่างกันอย่างมีนัยสำคัญ นอกนั้นไม่มีคู่ใดแตกต่างกันอย่างมีนัยสำคัญ

การวิเคราะห์สหสัมพันธ์ (Correlation analysis)

ความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปร 2 ตัว สามารถวัดได้ด้วยดัชนี 2 ตัว คือ สัมประสิทธิ์สหสัมพันธ์ (coefficient of correlation) และสัมประสิทธิ์เชิงทำนาย (coefficient of determination)

สัมประสิทธิ์สหสัมพันธ์ (coefficient of correlation) ซึ่งเป็นดัชนีบอกความสัมพันธ์หรือความเกี่ยวข้องระหว่างตัวแปร 2 ตัว เช่น ความสัมพันธ์ระหว่างความสามารถในการอ่านกับความสามารถในการฟัง หรือผลสัมฤทธิ์ในการเรียนภาษาไทยกับเจตคติต่อการเรียนภาษาไทย เป็นต้น

สถิติที่ใช้บอกค่าสหสัมพันธ์มีหลายตัว แต่ในที่นี้จะกล่าวถึงสหสัมพันธ์เพียร์สัน (Pearson's product moment correlation coefficient) เท่านั้น เนื่องจากเป็นที่รู้จักและใช้กันบ่อยที่สุด

สำหรับสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (r) นั้น ข้อมูลที่เหมาะสมที่สุดคือ คะแนนที่มีช่วงเท่ากัน (interval scale) ค่าของ r อยู่ระหว่าง −1 ถึง +1 หรืออยู่ระหว่าง 0 กับ ±1 ค่า r = 0 หมายถึง ตัวแปรทั้งสองไม่มีความสัมพันธ์กัน เมื่อ r = +1 หรือ −1 แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์กันชนิดสมบูรณ์ การอ่านค่า r ต้องคำนึงถึง 2 เรื่อง คือ ทิศทาง และความเข้มข้นของความสัมพันธ์ เครื่องหมาย + แสดงว่า ความสัมพันธ์เป็นไปในทิศทางเดียวกัน เช่น เด็กที่มีความสามารถสูงในการฟังก็จะมีความสามารถสูงในการอ่านด้วย ส่วนเครื่องหมาย − แสดงความสัมพันธ์ในทิศทางตรงข้าม เช่น เด็กที่ได้คะแนนภาษาอังกฤษสูงมีแนวโน้มที่จะได้คะแนนต่ำในวิชาคณิตศาสตร์ ส่วนตัวเลขแสดงความเข้มของความสัมพันธ์ ตัวเลขที่สูงขึ้นแสดงว่าตัว

แปรทั้งสองมีความสัมพันธ์ใกล้ชิดยิ่งขึ้น ตัวเลขที่ต่ำลงมาใกล้ศูนย์แสดงถึงความสัมพันธ์กันน้อยลง

ตัวอย่างงานวิจัยที่สามารถใช้การวิเคราะห์สหสัมพันธ์ได้ คือ งานวิจัยที่สำรวจความเข้าใจในการอ่านภาษาอังกฤษ (ตัวแปร Y) กับเจตคติในการเรียนวิชาภาษาอังกฤษ (ตัวแปร X) ของนักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 320 คน ผู้วิจัยวัดตัวแปรทั้งสองของนักเรียนทั้ง 320 คน ดังนั้นนักเรียนแต่ละคนจะมีคะแนน 2 ตัว ผู้วิจัยจึงมีคะแนนทั้งหมด 320 คู่ ซึ่งสามารถนำมาคำนวณค่า r ได้ โดยอาจคำนวณโดยใช้สูตรที่มีในหนังสือสถิติทั่วไป หรือใช้โปรแกรมคอมพิวเตอร์ก็ได้

ผู้วิจัยควรระมัดระวังในการตีความหมายของสัมประสิทธิ์สหสัมพันธ์ (r) ค่า r ไม่ว่าจะสูงเพียงใดเป็นเพียงการแสดงถึงการที่ตัวแปร 2 ตัว แปรผันตามกันหรือตรงข้าม แต่ไม่ได้บอกว่าตัวแปรใดเป็นสาเหตุ ตัวแปรใดเป็นผล เช่น ถ้าผู้วิจัยคำนวณสัมประสิทธิ์สหสัมพันธ์ระหว่างความเข้าใจในการอ่านภาษาอังกฤษ (Y) และเจตคติในการเรียนวิชาภาษาอังกฤษ (X) ของนักเรียนมีค่าเท่ากับ .85 (r = .85) ผู้วิจัยทราบเพียงว่า ตัวแปรทั้งสองนี้ผันไปในทิศทางเดียวกันสูง แต่ไม่สามารถสรุปได้ว่าตัวแปรใดเป็นเหตุตัวแปรใดเป็นผล

การตัดสินใจว่าตัวแปร 2 ตัวที่ศึกษามีความสัมพันธ์กันอย่างมีนัยสำคัญหรือไม่ในกลุ่มประชากร (การทดสอบสมมุติฐานว่าตัวแปร 2 ตัวมีความสัมพันธ์กันจริงไหม) ก็ใช้วิธีเปิดตารางจากหนังสือสถิติ หรืออ่านค่า p จากผลที่พิมพ์จากคอมพิวเตอร์เช่นกัน

ค่า r เมื่อนำมายกกำลังสองได้เป็น r^2 ดังนี้เรียกว่า เรียกกันว่า สัมประสิทธิ์เชิงทำนาย (coefficient of determination) มีความหมายว่า จากความแปรปรวนทั้งหมดของตัวแปร Y ถ้าทราบคะแนนตัวแปร X จะสามารถอธิบายความแปรปรวนของ Y ได้เป็นสัดส่วนเท่าใด หรือกล่าวได้ว่า r^2 เป็นสัดส่วนของความแปรปรวน Y ที่สามารถอธิบายได้ด้วยตัวแปร X

ตัวอย่างข้างบนที่ r = .85 ดังนั้น r^2 = .7225 จึงมีความหมายว่า 72.25% ของความแปร ปรวนใน Y สามารถอธิบายได้ด้วยความแปรปรวนใน X หรือในกรณีตัวอย่างนี้หมายความว่า จากความแปรปรวนทั้งหมดของคะแนนความเข้าใจในการอ่านภาษาอังกฤษนั้น 72.25% สามารถอธิบายได้ด้วยความแปรปรวนของคะแนนเจตคติในการเรียนวิชาภาษาอังกฤษ ซึ่งจะเป็นประโยชน์ในการทำนายคะแนนตัวแปร Y จากการทราบค่าตัวแปร X ในการวิเคราะห์การถดถอย (regression analysis)

การทดสอบไค-สแควร์ (Chi-square test - χ^2)

ในการวิจัยจำนวนมาก ผู้วิจัยรวบรวมข้อมูลเพียงความถี่ ซึ่งอยู่ในระดับนามบัญญัติ (nominal scale) (เช่น นับจำนวนคน) และต้องการทดสอบสมมุติฐานเช่นกัน สถิติที่เหมาะสมในกรณีเช่นนี้ คือ การทดสอบไค-สแควร์ (χ^2-test) ข้อมูล(ความถี่)นำมาเขียนลงในตารางที่เรียกว่า contingency table ซึ่งอาจมี 2 มิติ(สำหรับตัวแปร 2 ตัว)หรือมากกว่าก็ได้

การทดสอบไค-สแควร์ ใช้กับสมมุติฐาน 3 อย่าง คือ

1. การทดสอบความเหมาะ (goodness of fit test) นั่นคือ การตรวจสอบว่า การแจกแจงความถี่ในกลุ่มตัวอย่างมาจากประชากรที่มีการแจกแจงความถี่ที่กำหนดหรือไม่

2. การทดสอบการเท่ากัน (homogeneity) ของสัดส่วน ในกรณีนี้ผู้วิจัยรวบรวมข้อมูลเกี่ยวกับกับตัวแปรบางตัว เช่น การเลือกเรียนวิชาภาษาอังกฤษจากกลุ่มตัวอย่างหลายกลุ่ม อาจเป็น 2 กลุ่ม เช่น เพศชาย-หญิง และต้องการตรวจสอบว่าสัดส่วนการเกิดสำหรับตัวแปรที่สนใจเท่ากันหรือไม่ ผู้วิจัยต้องการทดสอบว่านักเรียนชายและนักเรียนหญิงมาจากประชากรที่มีสัดส่วนการเลือกเรียนวิชาภาษาอังกฤษเท่ากันหรือไม่

3. การทดสอบความเป็นอิสระ (independence) วิธีนี้เป็นการตรวจสอบว่าตัวแปร 2 ตัว เป็น อิสระต่อกันหรือไม่ ในกลุ่มประชากร ถ้าไม่เป็นอิสระต่อกัน แสดงว่าตัวแปรเหล่านั้นมีความสัมพันธ์กัน

การใช้สถิติไค-สแควร์ เพื่อการทดสอบในข้อ 2 และ 3 ใช้กันบ่อย การคำนวณค่า χ^2ใช้สูตรเดียวกัน แต่แตกต่างกันที่การรวบรวมข้อมูล และสมมุติฐาน

การใช้สถิติไค-สแควร์ไม่มีเงื่อนไขหรือข้อตกลงเบื้องต้นเกี่ยวกับการกระจายของประชากร (distribution free) ดังเช่นสถิติที่กล่าวมาก่อนหน้านี้ แต่หน่วยต่าง ๆ ในกลุ่มตัวอย่างต้องมีความเป็นอิสระต่อกัน

ตัวอย่างการใช้สถิตไค-สแควร์ (χ^2-test)

1. สมมติผู้วิจัยสนใจการเลือกเรียนวิชาภาษาอังกฤษของนักเรียนชาย-หญิง และสำรวจข้อมูลจากนักเรียนชาย 30 คน และนักเรียนหญิง 55 คน ได้ดังนี้

การตัดสินใจ	ชาย	หญิง	รวม
จะเลือก	12	24	36
จะไม่เลือก	18	31	49
รวม	30	55	85

ผู้วิจัยต้องการทดสอบสมมติฐานว่า นักเรียนชายและนักเรียนหญิงจะเลือกเรียนวิชาภาษาอังกฤษด้วยสัดส่วนที่เท่ากันหรือไม่

การคำนวณค่า χ^2 อาจใช้สูตรจากตำราสถิติหรือใช้โปรแกรมคอมพิวเตอร์ช่วยก็ได้

ผลการคำนวณค่า χ^2 ได้เท่ากับ .1064 เมื่อเทียบกับค่า χ^2 ที่อ่านจากตาราง $_{.95}\chi^2_1$ = 3.841 (ดูตำราสถิติสำหรับรายละเอียด) และค่า χ^2 ที่คำนวณได้นี้น้อยกว่าค่าจากตา

ราง ดังนั้น จึงไม่ปฏิเสธ Ho นั่นคือ ข้อมูลที่มีอยู่ไม่สามารถใช้เป็นหลักฐานที่แสดงว่านักเรียนชาย-หญิง เลือกเรียนวิชาภาษาอังกฤษด้วยสัดส่วนที่แตกต่างกัน

2. ในงานวิจัยหนึ่งที่ได้กล่าวถึงแล้ว (Banpho, 2004) ผู้วิจัยสำรวจ tense ที่ใช้ในประโยคเริ่มเรื่องจากวารสารการแพทย์ ผลการแจกแจงความถี่เป็นดังนี้ (Banpho, 2004, p.80)

Tense	วารสาร 1	วารสาร 2
1. Present simple	44	73
2. Present perfect	17	12
3. Other	11	10

สมมุติฐานของการวิจัยในงานวิจัยเช่นนี้ควรเป็น "สัดส่วนการใช้ tense ต่าง ๆ ในวารสารทั้ง 2 ไม่เท่ากัน" และสมมุติฐานที่ทดสอบ คือ Ho จึงเป็น "สัดส่วนการใช้ tense ต่าง ๆ ในวารสารทั้ง 2 เท่ากัน" ในกรณีนี้การคำนวณค่า χ^2 ได้เท่ากับ 5.025 แต่ค่า χ^2 ที่อ่านได้จากตาราง คือ $_{.95}\chi^2_2 = 5.99$ ดังนั้น จึงไม่ปฏิเสธ Ho ด้วยระดับ $\alpha = .05$ และสรุปว่ายังไม่มีหลักฐานที่แสดงว่า สัดส่วนการใช้ tense ต่าง ๆ ในวารสารทั้งสองไม่เท่ากัน

ตัวอย่างที่กล่าวแล้วทั้งสองเป็นการทดสอบสมมุติฐานการเท่ากันของสัดส่วน ต่อไปเป็นตัวอย่างการทดสอบสมมุติฐานของความเป็นอิสระ

3. สมมติข้อมูลที่เป็นผลการสำรวจนักเรียน 120 คน เมื่อจำแนกตามตัวแปร 2 ตัว คือ เพศชาย-หญิง และความเห็นเกี่ยวกับการกำหนดให้ภาษาจีนเป็นภาษาต่างประเทศภาษาที่ 1 (เห็นด้วย ไม่เห็นด้วย ไม่แน่ใจ) ตารางการแจกแจงความถี่เป็นดังนี้

การตัดสินใจ	เห็นด้วย	ไม่เห็นด้วย	ไม่แน่ใจ	รวม
ชาย	29	36	15	80
หญิง	14	24	2	40
รวม	43	60	17	120

ผู้วิจัยต้องการตรวจสอบว่าเพศกับความคิดเห็นมีความเกี่ยวข้องกันหรือไม่

ในกรณีนี้ Ho คือ เพศกับความคิดเห็นไม่เกี่ยวข้องกัน

H_1 คือ เพศกับความคิดเห็นมีความเกี่ยวข้องกัน

ค่าสถิติ χ^2 ที่คำนวณได้มีค่าเท่ากับ 4.776 แต่เมื่ออ่านค่า χ^2 จากตารางแล้วพบว่า $_{.95}\chi^2_2 = 5.991$ ดังนั้น χ^2 ที่คำนวณได้มีค่าต่ำกว่า χ^2 ที่อ่านได้จากตาราง จึงไม่ปฏิเสธ Ho นั่นคือ ยังไม่มีหลักฐานยืนยันว่าตัวแปรทั้งสองมีความเกี่ยวข้องกัน

ตัวอย่างที่ 1 และ 2 เป็นการใช้ไค-สแควร์ในการทดสอบในกรณีข้อ 2 คือการเท่ากันของสัดส่วน ส่วนตัวอย่างที่ 3 เป็นการทดสอบในกรณีข้อ 3 คือการทดสอบความเป็นอิสระ เป้าหมายในการทดสอบสมมุติฐานแตกต่างกัน และการรวบรวมข้อมูลก็แตกต่างกันดังได้กล่าวแล้ว

สรุป ผู้วิจัยที่สนใจข้อมูลเชิงปริมาณอาจใช้สถิติบรรยายหรือสถิติอนุมาน สถิติบรรยายเพียงแต่บอกได้ว่ามีอะไร เท่าไร ส่วนสถิติอนุมานสามารถสรุปเกี่ยวกับลักษณะของประชากรโดยอาศัยข้อมูลจากกลุ่มตัวอย่าง การเลือกใช้สถิติที่เหมาะสมควรพิจารณาลักษณะของข้อมูลและคำถามการวิจัย

เอกสารอ้างอิง

เกศแก้ว เพ็งพริ้ง. 2547. *ความสามารถในการอ่านและการเขียนคำถูกอักขรวิธีของนักเรียนชั้นมัธยมศึกษาปีที่ 6 โรงเรียนสาธิตมหาวิทยาลัยศรีนครินทรวิโรฒ ปทุมวัน.* สารนิพนธ์ กศ.ม. (ภาษาไทย). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

นฤมล เทียบสวัสดิ์. 2545. *การเปรียบเทียบความเข้าใจในการอ่านและเจตคติในการเรียนภาษาอังกฤษของนักเรียนชั้นมัธยมศึกษาปีที่ 3 ที่เรียนการอ่านด้วยการฝึกใช้ความรู้เดิม โดยใช้เทคนิคโครงสร้างระดับยอดกับวิธีสอนอ่านตามคู่มือครู.* ปริญญานิพนธ์ กศ.ม. (การมัธยมศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

ผจงจิต อินทสุวรรณ. 2528. *สถิติอนุมาน.* สถาบันวิจัยพฤติกรรมศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ

พิณทิพย์ ทวยเจริญ. 2543. "สัมฤทธิ์ผลในการสอนภาษาอังกฤษให้แก่เด็กเริ่มเรียน", *วารสารภาษาและภาษาศาสตร์.* ปีที่ 19 ฉบับที่ 1 (กรกฎาคม-ธันวาคม): 9-31.

วรรณา ฐิติโรจน์ไพบูลย์. 2536. *การสังเคราะห์งานวิจัยทางภาษาศาสตร์เชิงจิตวิทยา.* ปริญญานิพนธ์ กศ.ม. (ภาษาศาสตร์การศึกษา). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

วัสสิกา รุมาคม. 2545. *ลักษณะภาษาโฆษณาสำหรับผู้หญิงในสื่อนิตยสารตั้งแต่ปี พ.ศ. 2475 ถึง พ.ศ. 2543.* วิทยานิพนธ์ ศศ.ม. (ภาษาศาสตร์). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยธรรมศาสตร์.

วิศนี เครือวนิชกรกุล. 2547. "การศึกษารูปแบบภาษานิยมในบทสนทนาภาษาไทย", *ภาษาและวัฒนธรรม.* ปีที่ 23 ฉบับที่ 1 (มกราคม-มิถุนายน): 44-56.

อัญชลี วรัญญาบริรักษ์. 2546. *การวิเคราะห์ท่วงท่าของการประพันธ์ในสุภาษิตพระร่วง.* ปริญญานิพนธ์ กศ.ม. (ภาษาไทย). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ.

Banpho, P. 2004. "A survey of writing techniques, authorship, tenses, voices, and length of opening sentences of medical research articles", *PASSA.* Vol. 35 (April): 69-87.

Field, A. 2000. *Discovering Statistics Using SPSS for Windows.* London: SAGE Publications.

McCarthey, S. J. ; Guo, Y. ; & Cummins, S. 2005. "Understanding changes in elementary mandarin students' L1 and L2 writing", *Journal of Second Languages Writing*. Vol. 14, (2) June: p. 71-104.

Wilcox, R. R. 1996. *Statistics for the Social Sciences*. San Diego, CA: Academic Press.