งานวิจัยฉบับสมบูรณ์

ความเสี่ยงของการเกิดความคาดเคลื่อนในการวิจัยเชิงปริมาณด้านสังคมวิทยา

Risk of Error in the Quantitative Sociology Research

เรวัต แสงสุริยงค์

งานวิจัยนี้ได้รับทุนอุดหนุนจากงบประมาณเงินรายได้
ประจำปีงบประมาณ 2561
ภาควิชาสังคมวิทยา คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยบูรพา

คำนำ

การวิจัยเป็นเครื่องมือประเภทหนึ่งที่นิยมนำไปใช้ทั้งในการค้นหาความรู้และสนับสนุนความรู้ แม้ว่าจะมีทั้งวิธีเชิงปริมาณและเชิงคุณภาพ แต่การวิจัยทั้งสองแบบมีระเบียบวิธีที่ได้รับการยอมรับกันอย่างแพร่หลายว่าจะทำให้ได้คำตอบที่ถูกต้อง แต่ดูเหมือนว่าการละเมิดข้อตกลงที่กำหนดไว้ในแต่ละขั้นตอนของการดำเนินการวิจัยแต่ละวิธีเกิดขึ้นได้เสมอทั้งด้วยความตั้งใจและไม่ตั้งใจ และด้วยความรู้และความไม่รู้

อาจกล่าวได้ว่า การค้นหาความรู้ด้านสังคมวิทยาตกอยู่ภายใต้อิทธิพลของการค้นหาความรู้ที่เป็นวิทยาศาสตร์มาตั้งแต่เริ่มก่อตัว จึงมีการใช้วิจัยเชิงปริมาณมาอย่างต่อเนื่อง รวมถึงการถ่ายทอดความรู้ด้านสังคมวิทยาในระดับบัณฑิตศึกษาหรือสูงกว่าปริญญาตรีในประเทศไทย และมีการใช้วิธีการวิจัยเชิงปริมาณในการทำวิทยานิพนธ์และดุษฎีนิพนธ์กันอย่างแพร่หลาย

ด้วยความสงสัยว่า การทำวิทยานิพนธ์และดุษฎีนิพนธ์ที่ใช้วิธีเชิงปริมาณมีการดำเนินตามระเบียบแบบแผนและข้อตกลงที่เชื่อว่าจะทำให้ได้ผลการวิจัยที่ถูกต้องหรือไม่ จึงทำให้เกิดความสนใจที่จะค้นหาคำตอบว่าการทำวิทยานิพนธ์และดุษฎีนิพนธ์มีการละเมิดข้อตกลงที่อาจส่งผลให้เกิดความคลาดเคลื่อนของผลการศึกษามากน้อยเพียงใด รวมถึงการค้นหาปัจจัยที่มีผลต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย เพื่อเป็นแนวทางในการตรวจสอบและประเมินความเสี่ยงการเกิดความคลาดเคลื่อนในการวิจัย โดยได้รับการสนับสนุนงบประมาณจากภาควิชาสังคมวิทยา คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยบูรพา และได้รับการรับรองผลการพิจารณาจริยธรรมการวิจัยในมนุษย์ จากมหาวิทยาลัยบูรพา ภายใต้รหัสโครงการวิจัย Hu 001/2561

เรวัต แสงสุริยงค์

กุมภาพันธ์ 2563

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์ในการสำรวจการออกแบบการวิจัยตามระเบียบวิธีการวิจัยเชิงปริมาณและวิเคราะห์หาความเสี่ยงของการเกิดความคลาดเคลื่อนของผลการวิจัย โดยใช้วิธีการเชิงปริมาณรวบรวมข้อมูลจากกลุ่มตัวอย่างที่เป็นวิทยานิพนธ์และดุษฎีนิพนธ์ในฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทยของสำนักงานคณะกรรมการการอุดมศึกษา

กลุ่มตัวอย่างเป็นงานวิจัยเชิงปริมาณสาขาสังคมวิทยาที่ได้จากการรวบรวมข้อมูลด้วยโปรแกรมค้นหาข้อมูล การแปลงข้อมูลด้วยภาษา XML (eXtensible Markup Language) และสำรวจข้อมูลด้วยแบบสำรวจดิจิทัล

การวิเคราะห์จัดกลุ่มแบบเคมีน (K-means Cluster Analysis) แบ่งกลุ่มงานวิจัยได้จำนวน 30 กลุ่ม ที่มีวิธีการดำเนินการวิจัยที่เหมือนกันในด้านตัวแปรสมมติฐาน สถิติ เครื่องมือวัด ระดับการวัด การตรวจสอบข้อมูล ประชากร ขนาดตัวอย่าง และการเลือกตัวอย่าง และในจำนวน 30 กลุ่ม มีงานวิจัยที่มีการทดสอบสมมติฐานจำนวน 26 กลุ่ม ที่นำมาทดสอบเชิงสถิติระหว่างตัวอย่างงานวิจัยที่มีกับไม่มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยพบว่า มีจำนวนสัดส่วนที่ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ (Sig. = 0.286)

การวิเคราะห์การถดถอยแบบโลจิสติคพหุ (Multiple Logistic Regression: MLR) โดยใช้วิธีนำตัวแปรทำนายเข้าวิเคราะห์ในแบบจำลองพร้อมกันทั้งหมด (Enter Method) พบว่า สมการที่ประกอบด้วยตัวแปรทำนายเครื่องมือวัด ระดับการวัด ประชากร ขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล ไม่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ (Sig. = 0.999) แต่จากผลการวิเคราะห์โดยใช้วิธีคัดเลือกตัวแปรทำนายเข้าแบบจำลองอย่างเป็นขั้นตอน (Forward Stepwise) ได้สมการที่ประกอบด้วยตัวแปรทำนายเครื่องมือวัด และขนาดตัวอย่าง ที่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ (Sig. = 0.000) โดยตัวแปรทำนายสามารถทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยได้ประมาณร้อยละ 34.2 (Nagelkerke R^2 = 0.342) และแบบจำลองมีประสิทธิภาพในการทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยได้ประมาณร้อยละ 68.9 และสามารถทำนายได้ว่างานวิจัยที่มีการละเมิดข้อตกลง คือ ไม่มีการตรวจสอบเครื่องมือวัดและไม่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง มีความเสี่ยงในการเกิดความคลาดเคลื่อนในงานวิจัย (P(y) = 0.93) โดยงานวิจัยที่ไม่มีการทดสอบเครื่องมือวัด มีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมากกว่างานวิจัยที่มีการทดสอบเครื่องมือวัดประมาณ 14 เท่า (Odds Ratio = 14.059) และงานวิจัยที่ไม่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง มีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมากกว่างานวิจัยที่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่างประมาณ 2 เท่า (Odds Ratio = 2.322)

Abstract

The research was aimed to explore research designs available in quantitative research methodology and to analyze risk in error occurred in research results, by using quantitative methods to collect samples of theses and dissertations in sociology available in ThaiLIS-Thai Library Integrated System of the Higher Education Commission Office.

The samples were quantitative research projects in sociology drawn by search engine, eXtensible Markup Language (XML), and digital questionnaire.

The K-means Cluster Analysis was applied for classifying the samples into 30 clusters, based on variables including hypothesis, statistics, instruments, scale of measurement, data exploratory, population, sample sizes and sampling. Among 30 clusters, only 26 groups with tested hypothesis had been chosen in the sample for further analysis. The results of statistical testing between the hypothesis tested research groups with and without risk of error in research had no significantly statistical difference in ratio (Sig. = 0.286).

Through Multiple Logistic Regression (MLR) Analysis by Enter Method, in the formula that comprised all predicting variables including instruments, scale of measurement, data exploratory, population, sample sizes, and sampling, it had no co-affect on risk of error in research (Sig. = 0.999). However, from result of analysis by selecting predicting variables in Forward Stepwise Method, it resulted in a formula consisted of predicting measurement variables, and sample size. It had co-affect on risk of error in research (Sig. = 0.000). The predictor variables could predict risk of error in research around 34.2 percent (Nagelkerke R^2 = 0.342). The model had 68.9 percent predicted the correction of classification of risk of error in research.

It could be predicted that researches with violated assumption, including unvalidated instrument and non statistic calculation of sample size, were likely to have risk of error (P(y) = 0.93). The researches with unvalidated instruments had 14 times higher risk of error than the researches with validated instruments (Odds Ratio = 14.059). The researches without statistical calculation for sample size also had two times higher risk of error than the researches with statistical calculation for sample size (Odds Ratio = 2.322).

สารบัญ

สารบัญ (ต่อ)

สารบัญ (ต่อ)

สารบัญตาราง

สารบัญตาราง (ต่อ)

สารบัญแผนภูมิ

สารบัญภาพ

บทที่ 1

บทนำ

ความเป็นมาของปัญหาการวิจัย

ปรากฏการณ์ทางสังคมในยุโรปที่เกิดขึ้นช่วงศตวรรษที่ 18 คือ ความยุ่งเหยิงทางการเมืองที่ทำให้เกิดการปฏิวัติในฝรั่งเศส และปัญหาสังคมที่เกิดจากการปฏิวัติอุตสาหกรรมในอังกฤษ ไม่สามารถอธิบายและให้คำตอบได้ด้วยความคิดทางสังคมที่ตกอยู่ภายใต้อิทธิพลของลัทธิเทววิทยา (theological) และระบบศักดินา (feudal) ที่มีอยู่ในขณะนั้น ทำให้มีนักปรัชญากลุ่มหนึ่งพยายามลบล้างอิทธิพลดังกล่าวและนำเอาวิธีการศึกแบบวิทยาศาสตร์ธรรมชาติ (natural sciences) มาใช้ในการศึกษาสังคม

เฮนรี เดอร์ เซ็นต์-ไซมอน (Comte Henri de Saint-Simon) นักปรัชญาชาวฝรั่งเศสได้เสนอให้ใช้วิธีการศึกษาแบบวิทยาศาสตร์ธรรมชาติมาใช้ศึกษาและวิเคราะห์สังคม เพื่อนำความรู้ไปใช้ในการปฏิรูปสังคม (social reform) และพยากรณ์แผนการที่จะต้องกระทำในสังคมอุตสาหกรรม ซึ่งออกุสต์ กองต์ (Auguste Comte) ที่เป็นทั้งลูกศิษย์และเพื่อนร่วมงานก็เห็นด้วยกับข้อเสนอดังกล่าว และยอมรับว่าวิธีการดังกล่าวมีการพัฒนาและใช้กันอยู่แล้วในสาขาวิชาวิทยาศาสตร์ธรรมชาติ

การศึกษาด้วยหลักปรัชญาของความเป็นจริงที่เรียกว่า ปฏิฐานนิยม (positive philosophy) จะทำให้ได้ความรู้ทางสังคมที่อยู่บนฐานของความเป็นจริง (true) ปราศจากอิทธิพลของเหตุผลที่อยู่เหนือธรรมชาติ (metaphysical) อุดมคติหรืออุดมการณ์ (ideological) ศาสนา (religious) หรือค่านิยมของสังคม (moral values) โดยงานเขียนช่วงแรกของกองต์ เรียกศาสตร์ใหม่ที่ใช้ในการศึกษาสังคมว่า การศึกษาธรรมชาติของปรากฏการณ์ทางสังคม (social physics) จนกระทั่งกองต์พบว่าคำดังกล่าวมีนักสถิติชาวเบลเยี่ยมใช้อยู่แล้ว ในปี ค.ศ. 1837 กองต์จึงเริ่มนำเอาคำว่า สังคมวิทยา (sociology) ที่เอ็มมานูเอล โฌแซฟ ซีแยส์ (Emmanuel-Joseph Sieyes) เคยเขียนไว้ในงานที่ไม่ได้ตีพิมพ์ในปี ค.ศ. 1780 มาใช้แทนตั้งแต่นั้นเป็นต้นมา

สังคมวิทยามีต้นกำเนิดอยู่ในทวีปยุโรปในช่วงศตวรรษที่ 19 ต่อจากนั้นไม่นานประมาณศตวรรษที่ 20 แนวคิดเกี่ยวกับสังคมวิทยาได้แผ่ขยายเข้าสู่ทวีปอเมริกาเหนือที่ประเทศสหรัฐอเมริกาเป็นประเทศแรก โดย

มหาวิทยาลัยเยล (Yale University) เปิดสอนวิชาสังคมวิทยาเป็นแห่งแรกในปี ค.ศ. 1876 และมหาวิทยาลัยชิคาโก (Chicago University) เปิดภาควิชาสังคมวิทยาเป็นแห่งแรกในปี ค.ศ. 1892 และสหรัฐอเมริกาได้กลายมาเป็นประเทศมีความเจริญรุ่งเรืองด้านสังคมวิทยาเป็นอย่างมาก

เมื่อนักวิชาการจากประเทศไทยเดินทางไปศึกษาในประเทศสหรัฐอเมริกาทำให้มีการนำเอาแนวคิดทางสังคมวิทยาเข้ามาเผยแพร่ในวงการการศึกษาในประเทศไทย โดยเริ่มสอนที่โรงเรียนข้าราชการฝ่ายพลเรือนของกระทรวงมหาดไทยในปี พ.ศ. 2473 ต่อมาเริ่มสอนในชื่อวิชาอาชญาวิทยาในระดับอุดมศึกษาที่มหาวิทยาลัยธรรมศาสตร์และการเมืองในปี พ.ศ. 2478 และมีการสอนอย่างเป็นระบบในปี พ.ศ. 2491 โดยศาสตราจารย์ ดร. เกษม อุทยานิน ที่ศึกษาวิชานี้โดยตรง และต่อมาได้ขยายตัวจากมหาวิทยาลัยในกรุงเทพมหานครออกไปมหาวิทยาลัยในส่วนภูมิภาค (สัญญา สัญญาวิวัฒน์, 2525; มหาวิทยาลัยเชียงใหม่, 2530; เรวัต แสงสุริยงค์, 2541)

การเปิดสอนหลักสูตรระดับบัณฑิตศึกษาหรือสูงกว่าปริญญาตรีในมหาวิทยาลัย เช่น หลักสูตรระดับปริญญาโทและปริญญาเอก ที่กำหนดให้ผู้เรียนต้องทำวิทยานิพนธ์หรือดุษฎีนิพนธ์ (theses or dissertations) เพื่อแสดงให้เห็นว่า ผู้เรียนมีความรู้และความเข้าใจในแนวคิดและทฤษฎี สามารถนำไปประยุกต์ใช้ในเชิงวิชาการและแก้ไขปัญหาสังคม รวมถึงการพัฒนาต่อยอดแนวคิดและทฤษฎีด้วยการหาความรู้ใหม่ที่เป็นประโยชน์เพื่อเผยแพร่ต่อวงการวิชาการและสังคม (contributions) โดยมีระเบียบวิธีการวิจัยเป็นเครื่องมือที่สำคัญในการทำวิทยานิพนธ์หรือดุษฎีนิพนธ์ เพราะเป็นกระบวนการที่มีวิธีการและขั้นตอนที่สามารถตรวจสอบได้ (verified) และวงการวิชาการต่างยอมรับว่าความรู้และคำตอบที่ได้จากการวิจัยมีความเที่ยงตรง (validity) และความน่าเชื่อถือ (reliability)

นักสังคมวิทยาจำนวนมากเดินตามแนวคิดในการศึกษาสังคมของนักสังคมวิทยารุ่นบุกเบิก นิยมใช้เทคนิคการศึกษาเหมือนกับการศึกษาด้านวิทยาศาสตร์ธรรมชาติหรือที่เรียกว่า วิธีการที่เป็นวิทยาศาสตร์ (scientific method) อย่างไรก็ตามการศึกษาสังคมมนุษย์ด้วยวิธีการที่เป็นวิทยาศาสตร์เพียงอย่างเดียวไม่สามารถที่จะอธิบายปรากฏการณ์ต่าง ๆ ที่เกิดขึ้นในสังคมมนุษย์ได้อย่างเพียงพอ แมค เวเบอร์ (Max Weber) นักสังคมวิทยาชาวเยอรมันจึงเสนอวิธีการที่จะเข้าใจพฤติกรรมของมนุษย์โดยตรงที่เรียกว่า “ Sympathetic

Understanding” ซึ่งเป็นวิธีการศึกษาสังคมมนุษย์ด้วยการเข้าไปอยู่ภายในเหตุการณ์หรือเรื่องที่ต้องการจะศึกษา

แม้ว่าจะมีข้อถกเถียงเกี่ยวกับภาววิทยา (ontology) และญาณวิทยา (epistemology) ในการศึกษาพฤติกรรมของมนุษย์และปรากฏการณ์ทางสังคมในเชิงสังคมวิทยาอย่างกว้างขวางและต่อเนื่อง แต่สามารถแบ่งระเบียบวิธีวิทยา (methodology) ในการศึกษาและหาความรู้ในสาขาวิชาสังคมวิทยาได้เป็น 2 วิธีหลัก คือ วิธีเชิงปริมาณ (quantitative method) และวิธีเชิงคุณภาพ (qualitative method)

การได้มาซึ่งข้อมูลและการวิเคราะห์ข้อมูลเพื่อตอบปัญหาการวิจัยแต่ละวิธีต่างมีกระบวนการและวิธีการให้ได้มาซึ่งคำตอบที่ถูกต้อง สำหรับการวิจัยเชิงปริมาณที่ใช้วิธีการทางสถิติ (statistical) เป็นเครื่องมือในการสรุปผลการวิจัยเพื่อยอมรับ (accept) หรือปฏิเสธ (reject) สมมุติฐาน (hypothesis) ที่ได้มาจากแนวคิด (concept) หรือทฤษฎี (theory) ตามหลักนิรนัย (deduction) เป็นการวิจัยที่ผู้เรียนในหลักสูตรระดับปริญญาโทและปริญญาเอกนิยมนำมาใช้ทำวิทยานิพนธ์และดุษฎีนิพนธ์กันอย่างแพร่หลาย

ผู้วิจัยที่เลือกใช้วิธีการวิจัยเชิงปริมาณนอกจากมีความรู้ในระเบียบวิธีการวิจัยและการสร้างเครื่องมือในการรวบรวมข้อมูลเป็นอย่างดีแล้ว ยังต้องมีความรู้และความเข้าใจในการใช้สถิติเป็นอย่างดี รวมถึงต้องนำเอาข้อมูลมาใช้ให้ถูกต้องตามข้อตกลงเบื้องต้น (basic assumptions) ของแต่ละสถิติด้วย เพื่อป้องกันมิให้เกิดผลการวิเคราะห์เกิดความคลาดเคลื่อนในการอ้างอิง (inferential errors) ไปสู่แนวคิดหรือทฤษฎีที่ต้องการพิสูจน์หรือตรวจสอบ

ความคลาดเคลื่อนในการวิจัยมีความเกี่ยวข้องตั้งแต่เริ่มทำจนถึงสรุปผล การวิจัยแต่ละวิธีแม้ว่าจะมีเครื่องมือที่มีทั้งความเที่ยงตรงและความเชื่อมั่น แต่หากดำเนินการตามกระบวนการและขั้นตอนอย่างถูกต้องก็อาจได้ข้อสรุปที่ไม่ถูกต้องก็ได้ โดยเฉพาะการศึกษาที่ใช้ข้อมูลขนาดใหญ่มีโอกาสมากที่จะเกิดความคลาดเคลื่อนอย่างเป็นระบบ (systematic errors) จากข้อมูลที่ได้มา (Lash & Others, 2016) ทำให้นักวิจัยในหลายสาขาวิชาหันมาให้ความสำคัญกับความคลาดเคลื่อนในการวัด (measurement errors) กันมากขึ้น ส่วนในการวิจัยด้านสังคมวิทยามีการโต้เถียงกันไม่มากนักเกี่ยวกับผลกระทบที่เกิดจากการวิจัย (Thye, 2000)

การใช้เทคนิคการอ้างอิงเชิงสถิติจะมีความแม่นยำ (precise) และเชื่อได้ (rely) หรือไม่ มีความเกี่ยวข้องกันระหว่างความคลาดเคลื่อนของการเลือกตัวอย่าง (sampling error) ขนาดของกลุ่มตัวอย่าง

(sample size) กลุ่มตัวอย่างมีขนาดใหญ่มากพอหรือมากกว่า 30 หน่วยตัวอย่าง (Central Limit Theorem: CLT) และข้อมูลต้องได้มาจากตัวอย่างแบบสุ่ม (random sample) แต่อย่าคิดว่าเป็นเรื่องง่ายและเป็นคำตอบที่สมบูรณ์แล้ว (Neuman, 2006)

การศึกษาเกี่ยวกับความคลาดเคลื่อนมี 2 แนวคิด คือ การศึกษาในเชิงสถิติโดยใช้ทฤษฎีการเลือกตัวอย่างเชิงสถิติ (statistical sampling theory) และการศึกษาในเชิงการวัดโดยใช้ทฤษฎีการวัด (measurement theory) แต่ละแนวคิดระบุที่มาของความคลาดเคลื่อนไว้ ดังนี้ (Biemer & others, 1991: 1, 113-123; Holmes, 2004)

1. ความคลาดเคลื่อนเชิงสถิติ มีทั้งในลักษณะของความลำเอียง (bias) และความไม่แม่นยำ (imprecision) โดยความลำเอียงนั้นเป็นความคลาดเคลื่อน (error) ที่มีลักณะความสอดคล้องกัน (consistent tendency) และเป็นไปในทิศทางเดียวกัน (direction) ส่วนความไม่แม่นยำเป็นความคลาดเคลื่อนที่ไม่สามารถระบุสาเหตุและทิศทางได้ (nondirectional noise) โดยแบ่งสาเหตุได้เป็น 5 ประเภท คือ การเลือกตัวอย่าง (sampling) การวัด (measurement) การประมาณการ (estimation) การทดสอบสมมติฐาน (hypothesis testing) และการรายงาน (reporting)

2. ความคลาดเคลื่อนเชิงการวัด หรือที่เรียกว่า ความคลาดเคลื่อนแบบที่ 3 (type III error) สาเหตุที่ทำให้การสำรวจได้ข้อมูลไม่ถูกต้องในการวัดทุกประเภทเกิดมาจากทั้งความคลาดเคลื่อนในกระบวนการวัดที่ไม่ถูกต้อง และความลำเอียงที่ทำให้ได้คำตอบที่แตกต่างไปจากค่าที่เป็นจริง โดยจำแนกได้เป็น 3 สาเหตุ คือ การวัดตัวแปรผิด (wrong variable) การวัดไม่ถูกต้องหรือลำเอียง และเครื่องมือวัดมีความแปรปรวน (variability) มาก

ความสำคัญของปัญหาการวิจัย

ความคลาดเคลื่อนเชิงสถิติ (statistical errors) เป็นสิ่งสำคัญในการวิจัยที่เป็นวิทยาศาสตร์ (scientific research) และเห็นได้เสมอในการทบทวนงานวิจัย (Ghasemi & Zahediasl, 2012: 486) ความคลาดเคลื่อนเชิงอ้างอิง (inferential errors) และอำนาจหรือกำลังการจำแนกทางสถิติ (power) เป็นข้อมูลที่สำคัญในการประเมินเชิงปริมาณเพื่อใช้ในการวางแผนการวิจัยเชิงสำรวจที่ใช้วิธีการทางสถิติ และใน

การประเมินเพื่อกำหนดความน่าจะเป็นของความถูกต้องและความไม่ถูกต้องในการอ้างอิงสำหรับการวางแผนการวิเคราะห์ จึงต้องทำการออกแบบการตรวจสอบเพราะข้อมูลทั้งมีและไม่มีผลกระทบ นักวิจัยต้องระวังระหว่างความคลาดเคลื่อนแบบที่ 1 (type I error) กับความคลาดเคลื่อนแบบที่ 2 (type II error) และความสัมพันธ์ระหว่างความคลาดเคลื่อนแบบที่ 2 กับอำนาจหรือกำลังของสถิติ (statistical power) เพราะความคลาดเคลื่อนดังกล่าวมีผลกระทบอย่างสำคัญในการวางแผนการวิจัยที่เป็นวิทยาศาสตร์ (Rigby, 1998)

การศึกษาเกี่ยวกับความคลาดเคลื่อนในการวิจัยด้านสังคมศาสตร์ยุคบุกเบิกมีหลักฐานปรากฏให้เห็นในปี ค.ศ. 1926 จากงานของสจ๊วต เอ ไรซ์ (Struart A. Rice) เรื่อง “Contagious bias in the interview” ที่ตีพิมพ์ในวารสารอเมริกันด้านสังคมวิทยา (American Journal of Sociology) ที่วิเคราะห์ให้เห็นความลำเอียงในการตอบคำถามของผู้ให้ข้อมูล (informant) ที่ผันแปรไปตามการพูดคุยในกระบวนการสัมภาษณ์จากลักษณะบุคลิกภาพของผู้สัมภาษณ์ (Rice, 1929; Deming, 1944)

ในปี ค.ศ. 1979 ฮูเบิร์ต บลาล็อค (Hubert Blalock) นายกสมาคมสังคมวิทยาอเมริกัน (American Sociological Association) ได้กล่าวว่า การวัด (measurement) เป็นปัญหาสำคัญและร้ายแรงที่กำลังเผชิญอยู่ในการเรียนการสอนด้านสังคมวิทยา แต่หลังจากนั้นเป็นต้นมามีการเปลี่ยนแปลงไม่มากนัก ทุกวันนี้นักวิจัยยังคงพบกับปัญหาในการวัดและการสร้างกรอบแนวคิด (conceptualization) ที่ไม่แตกต่างไปจากที่บลาล็อคกังวลไว้ตั้งแต่ในอดีต ความคลาดเคลื่อนเชิงสถิติ (statistical error) ในงานวิจัยเชิงปริมาณของนักสังคมศาสตร์คงเป็นเรื่องซ้ำซากเช่นเดียวกับงานวิจัยเชิงวิทยาศาสตร์ของนักวิทยาศาสตร์ที่เกิดขึ้นมาอย่างยาวนาน เพราะการขาดความรู้และความเข้าใจเกี่ยวกับพื้นฐานด้านสถิติ (basic of statistics) ขาดการสอนที่ดีโดยนักสถิติขาดความสนใจและความเอาใจใส่หลังจากเรียนจบไปแล้ว จึงทำให้มีทักษะด้านสถิติที่ไม่เพียงพอในการทำวิจัยที่ต้องใช้สถิติ (Rigby, 1998; Thye, 2000) และมีการประมาณการว่าร้อยละ 50 ของบทความที่เป็นเชิงวิทยาศาสตร์ที่ตีพิมพ์ มีอย่างน้อย 1 บทความที่มีความคลาดเคลื่อนเชิงสถิติ (Curran-Everett & Benos, 2004)

กระบวนการตัดสินใจ (decision process) ในการทดสอบสมมติฐาน (hypothesis test) ในการวิจัยเชิงยืนยันเพื่อยอมรับ (accept) หรือเพื่อปฏิเสธ (reject) สมมติฐานหลัก (null hypothesis) ถ้าสมมติฐานหลักถูกปฏิเสธจะทำให้ตีความได้ว่าหลักฐานหรือข้อมูลที่นำมาใช้ในการทดสอบสนับสนุนสมมติฐานทางเลือก

(alternative hypothesis) หรือสมมติฐานเชิงทดลอง (experimental hypothesis) ดังนั้นผลการทดสอบหรือการทดลองถึงมีความเป็นไปได้สองผลลัพธ์ (two possible outcomes) สมมติฐานหลักอาจได้รับการยอมรับหรือปฏิเสธ หากกระบวนการตัดสินใจออกแบบการวิจัยเชิงยืนยันไม่ดีหรือเป็นการวิเคราะห์เชิงสำรวจทั่วไปจะทำให้เกิดอำนาจจำแนกทางสถิติต่ำ (underpowered) ไม่สามารถยอมรับหรือสนับสนุนสมมติฐานหลัก เพราะผลลัพธ์จากการวิจัยที่ไม่มีนัยสำคัญทางสถิติ (nonsignificant) สามารถทำให้อำนาจจำแนกต่ำกว่าที่จะทำให้สมมติฐานมีความเป็นจริง การวิเคราะห์อำนาจจำแนกทางสถิติตามแบบแผนหรือแบบดั้งเดิม (classical power analysis) ต้องกำหนดขนาดตัวอย่าง (sample size) ให้มีความน่าจะเป็นที่สูง (high probability) เพื่อให้ได้ข้อสรุปที่ถูกต้อง การวิเคราะห์อำนาจจำแนกทางสถิติต้องพิจารณาความน่าจะเป็นในการปฏิเสธของสมมติฐานหลักเมื่อสมมติฐานทางเลือกเป็นจริงและความน่าจะเป็นในการยอมรับสมมติฐานหลักเมื่อมีความเป็นจริง (Kennedy, 2015)

ดังนั้น ผลการวิจัยไม่เพียงแต่จะทำให้เกิดการปฏิเสธแนวคิดหรือทฤษฎีที่ได้รับการตรวจสอบและยืนยันจากข้อมูลเชิงประจักษ์ว่ามีความถูกต้องแล้วเท่านั้น แต่ยังทำให้เกิดข้อค้นพบใหม่ที่เชื่อถือไม่ได้ รวมถึงการนำเอาผลการวิจัยที่มีความคลาดเคลื่อนไปเผยแพร่และนำเข้าสู่ระบบฐานข้อมูลงานวิจัยและบทความวิชาการ ทำให้มีการนำไปใช้ในการอ้างอิงทั้งในเชิงวิชาการและการแก้ไขปัญหาสังคมตามมาอย่างต่อเนื่อง การพัฒนาความรู้และต่อยอดความรู้ด้วยการทำการวิจัยพื้นฐาน (basic research) และการวิจัยประยุกต์ (applied research) สะสมผลการวิจัยที่มีความคลาดเคลื่อนในการวัด (measurement error) ทำให้เกิดความหายนะอย่างต่อเนื่อง (Thye, 2000)

ปัจจุบัน วิทยานิพนธ์และดุษฎีนิพนธ์ของผู้จบหลักสูตรระดับบัณฑิตศึกษา (ปริญญาโทหรือปริญญาเอก) ที่เก็บไว้ในฐานข้อมูลของโครงการเครือข่ายห้องสมุดในประเทศไทย (ThaiLIS-Thai Library Integrated System) ของสำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ เพื่อเผยแพร่และให้บริการค้นคว้าสำหรับนำไปใช้ในการอ้างอิงทั้งเชิงวิชาการและการบริหาร ได้ผ่านการควบคุมโดยอาจารย์ที่ปรึกษาและกรรมการที่เป็นผู้เชี่ยวชาญด้านการวิจัยในแต่ละสาขาวิชามาแล้ว แต่ในการวิจัยเชิงปริมาณอาจมีการละเมิดข้อตกลงเบื้องต้น (basic assumptions) ในขั้นการออกแบบการวิจัยและการวิเคราะห์ข้อมูลทำให้เกิดความเสี่ยงในการเกิดความคลาดเคลื่อนแบบที่ 1 (type I error) กับความคลาดเคลื่อนแบบที่ 2 (type II error)

การวิจัยเพื่อค้นหาคำตอบอย่างเป็นระบบเพื่อศึกษาว่า วิทยานิพนธ์และดุษฎีนิพนธ์ที่ใช้วิธีการเชิงปริมาณมีการดำเนินการตามข้อตกลงเบื้องต้นในการวิจัยเชิงปริมาณหรือไม่ ซึ่งมีความเกี่ยวข้องกับความคลาดเคลื่อนในการตัดสินใจยอมรับและปฏิเสธสมมติฐานหรือตอบปัญหาการวิจัย ทำให้เกิดผลกระทบต่อการนำผลการวิจัยไปใช้พัฒนาวิชาการและการพัฒนาหรือแก้ไขปัญหาสังคมต่อไป

ปัญหาการวิจัย

การวิจัยเชิงปริมาณสาขาสังคมวิทยาของมหาวิทยาลัยในประเทศไทยที่เผยแพร่อยู่ในฐานข้อมูลของโครงการเครือข่ายห้องสมุดในประเทศไทยมีการออกแบบวิธีการดำเนินการวิจัยอย่างไร มีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมากน้อยเพียงใด มีการดำเนินการตามข้อตกลงเบื้องต้นในการวิจัยเชิงปริมาณมากน้อยเพียงใด มีปัจจัยใดที่มีผลต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย และมีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยมากน้อยเพียงใด

วัตถุประสงค์ของการวิจัย

1. เพื่อสำรวจการออกแบบวิธีการดำเนินการวิจัยเชิงปริมาณด้านสังคมวิทยา
2. เพื่อวิเคราะห์เปรียบเทียบการวิจัยเชิงปริมาณด้านสังคมวิทยาที่มีความเสี่ยงและไม่มีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย
3. เพื่อวิเคราะห์หาปัจจัยที่มีผลต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยเชิงปริมาณด้านสังคมวิทยา
4. เพื่อทำนายการเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนการวิจัยเชิงปริมาณด้านสังคมวิทยา

ขอบเขตของการวิจัย

งานวิจัยนี้ศึกษาเฉพาะวิทยานิพนธ์และดุษฎีนิพนธ์ (thesis and dissertation) ของหลักสูตรปริญญาโทหรือปริญญาเอก สาขาวิชาสังคมวิทยา ที่อยู่ในฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทย (ThaiLIS-Thai Library Integrated System) ของสำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ ที่แล้วเสร็จระหว่างปี พ.ศ. 2512-2560 เท่านั้น

ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นแนวทางในการประเมินความเสี่ยงการเกิดความคลาดเคลื่อนของผลการวิจัยเชิงปริมาณที่นำไปใช้อ้างอิงและต่อยอดการทำวิจัย

2. มหาวิทยาลัยและหน่วยงานที่เกี่ยวข้องนำไปต่อยอดประเมินความเสี่ยงการเกิดความคลาดเคลื่อนของผลการวิจัยเชิงปริมาณที่อยู่ในภารกิจของแต่ละหน่วยงาน

3. เป็นต้นแบบในการประเมินความเสี่ยงการเกิดความคลาดเคลื่อนของผลการวิจัยเชิงปริมาณสาขาวิชาอื่นๆ ที่อยู่ในฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทย (ThaiLIS-Thai Library Integrated System) ของสำนักงานคณะกรรมการการอุดมศึกษา และฐานข้อมูลงานวิจัยของหน่วยงานอื่นๆ

บทที่ 2

แนวคิดและทฤษฎี

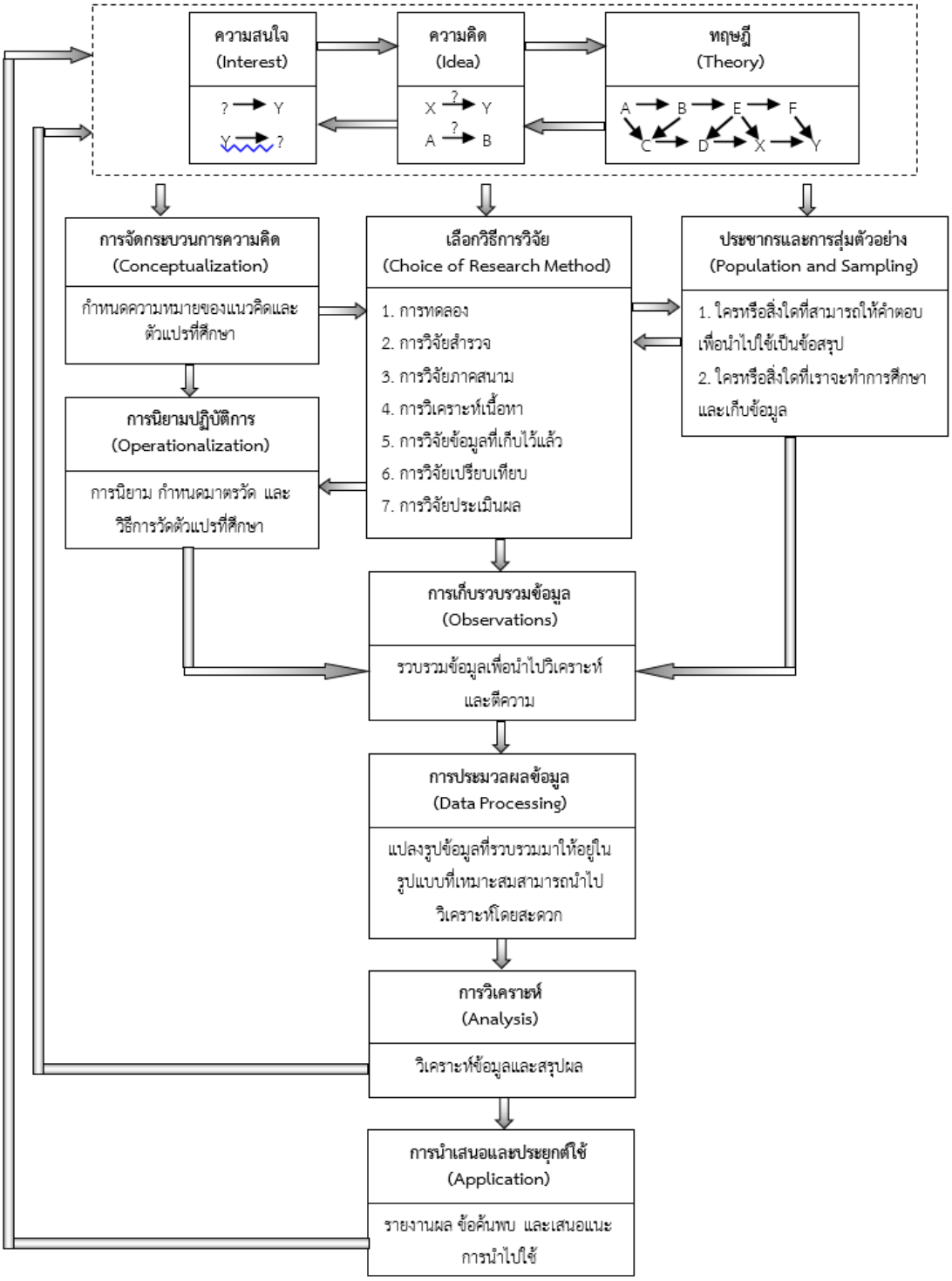
ความคลาดเคลื่อนในการวิจัยมีโอกาสเกิดขึ้นในทุกขั้นตอนในการทำวิจัย อาจเป็นความคลาดเคลื่อนที่มีผลกระทบอย่างเป็นระบบ (systematic) หรืออาจไม่เป็นระบบ (random) ตั้งแต่ขั้นการออกแบบการวิจัย (research design) จนถึงขั้นสรุปผลการวิจัย (conclusion) ดังนั้นในการทบทวนวรรณกรรมที่เกี่ยวข้องเพื่อสร้างกรอบแนวคิดจึงเน้นไปที่การวิจัยเชิงปริมาณ โดยมีรายละเอียดดังต่อไปนี้

การออกแบบการวิจัย

การวิจัยด้านสังคมศาสตร์จำแนกแบบกว้างๆ หรือตามลักษณะข้อมูลที่รวบรวมและเทคนิคการวิเคราะห์ได้เป็น 2 ประเภทใหญ่ คือ การวิจัยเชิงปริมาณ (quantitative research) และการวิจัยเชิงคุณภาพ (qualitative research) การวิจัยทั้งสองประเภทต่างมีความเชื่อเกี่ยวกับความจริงที่เรียกว่า ภาววิทยา (ontology) และวิธีการค้นหาความรู้หรือการเข้าถึงความจริงที่เรียกว่า ญาณวิทยา (epistemology) จึงทำให้เกิดมุมมองที่แตกต่างกันในวิธีการ (techniques) การรวบรวมข้อมูล (data collection) และการวิเคราะห์ข้อมูล (data analysis) (Blaikie, 2010: 204) แต่การวิจัยทั้ง 2 วิธี มีเป้าหมายที่เหมือนกัน คือ การได้ความรู้หรือข้อสรุปผลการวิจัยที่มีความถูกต้องและมีความน่าเชื่อถือ

การค้นหาความรู้ด้วยการวิจัยมีกระบวนการที่เป็นระบบ (systematic) การเริ่มทำการวิจัยจึงต้องมีการออกแบบการวิจัยอย่างเป็นขั้นตอน (step-by-step) การวิจัยทุกสาขาวิชาและทุกประเภทแม้ว่าจะมีฐานความเชื่อเกี่ยวกับภาววิทยาและญาณวิทยาที่แตกต่างกัน แต่มีขั้นตอนการทำวิจัยที่ไม่แตกต่างกันมากนัก (Scanlon cited by Wilkinson, 2000: 9)

ตามระเบียบวิธีของการทำวิจัย (research methodology) การวิจัยทางสังคมศาสตร์มีกระบวนการที่สามารถแสดงให้เห็นเป็นแผนผัง ตามภาพที่ 2.1



ภาพที่ 2.1 แสดงกระบวนการวิจัย (ปรับปรุงมาจาก Babbie, 2013: 113)

โดยทั่วไปยอมรับกันว่า ในทุกขั้นตอนของการทำวิจัยมีความลำเอียงหรืออคติ (bias) ของนักวิจัยเจือปนอยู่และยากที่จะทำให้ปราศจากความลำเอียง (objective) ทำให้ได้ข้อมูลที่มีความเบี่ยงเบน (deviance) ไปจากความจริง ส่งผลให้การวิเคราะห์เกิดความคลาดเคลื่อน (error) และได้ผลสรุปไม่ถูกต้อง จึงมีการออกแบบระเบียบวิธีการวิจัยให้นักวิจัยนำไปใช้ดำเนินการวิจัยอย่างเป็นระบบและเป็นขั้นตอน เพื่อป้องกันการเกิดความลำเอียงและความคลาดเคลื่อนให้มากที่สุด แต่ดูเหมือนว่ายังมีการละเมิดหรือไม่ปฏิบัติ (violated) ตามระเบียบวิธีการวิจัยอย่างเคร่งครัด

การวิจัยเชิงปริมาณ

ตั้งแต่ช่วงปลายศตวรรษที่ 19 เป็นต้นมา การวิจัยเชิงทดลองด้านจิตวิทยาเริ่มนำเอาวิธีการวิจัยเชิงปริมาณที่วางอยู่บนฐานความคิดแบบปฏิฐานนิยมมาใช้เป็นเครื่องมือค้นหาความรู้กันอย่างแพร่หลาย (Creswell, 2014: 12) ข้อมูลในการวิจัยเชิงปริมาณด้านสังคมส่วนใหญ่เมื่อเก็บรวบรวมมาแล้วนิยมทำให้เป็นรหัส (coding) ในรูปแบบของตัวเลข (numerical) สำหรับนำไปใช้กับโปรแกรมคอมพิวเตอร์เพื่อวิเคราะห์ด้วยวิธีการทางสถิติ (statistical analysis) ด้วยการพรรณนา (describing) และการอธิบาย (explaining) ก่อนนำไปนำเสนอตามจุดมุ่งหมาย (Babbie, 2013: 413-414)

การหาความรู้แบบนิรนัย (deductive approach) ด้วยวิธีการวิจัยเชิงปริมาณ โดยการนำเอาทฤษฎี (theories) ที่เป็นข้อสรุปที่เป็นนามธรรมจากความเป็นจริงมาทำให้เป็นรูปธรรม (operationalization) ด้วยการสร้างเป็นสมมติฐาน (hypothesis) ทำการนิยาม (identification) หรือพรรณนา (description) ตัวแปร (variable) สำหรับสร้างเป็นเครื่องมือ (instrument) วัดหรือรวบรวมข้อมูลที่เป็นตัวเลขจากกลุ่มตัวอย่างหรือประชากรที่มีจำนวนมากหรือขนาดใหญ่ เพื่อใช้ในการทดสอบหานัยสำคัญเชิงสถิติ (statistical significance) ของข้อมูล (Wagner & Okeke cited by Garner, M., Wagner, C., & Kawulich, B., 2009: 62; Williams, 2016: 88)

การวัดตัวแปรในด้านสังคมศาสตร์เป็นการกำหนดตัวเลขโดยอิงกับ (referring) คำตอบ (responses) ของบุคคลจากแบบสอบถาม (questionnaires) การกำหนดตัวเลขอาจเป็นเพียงการนับคำตอบแต่ละข้อคำถาม หรืออาจเป็นตัวเลขของคะแนน (number of point) รวมในการทดสอบความรู้หรือความสามารถ (Traub, 1994: 2-5)

เครื่องมือวัดและระดับการวัด

การวิจัยเชิงปริมาณมีฐานการวิเคราะห์มาจากแนวคิดแบบปฏิฐานนิยม (positivism) ภายใต้แนวคิดดังกล่าวเครื่องมือวัดต้องมีความเที่ยงตรง (validity) และความน่าเชื่อถือ (reliability) เพื่อให้ได้ข้อมูลที่ตรงกับความเป็นจริง เครื่องมือวัดจึงมีอิทธิพลต่อข้อมูล หากเกิดความคลาดเคลื่อนจากการวัด (measurement error) การวัดไม่สามารถวัดได้อย่างเที่ยงตรงก็จะทำให้เกิดความคลาดเคลื่อนแบบเป็นระบบ (systematic error) และไม่สามารถวัดได้อย่างน่าเชื่อถือก็จะทำให้เกิดความคลาดเคลื่อนแบบสุ่ม (random error) ทำให้ค่าเฉลี่ยของข้อมูลเพิ่มขึ้นหรือลดลงทำให้ได้ผลลัพธ์ที่เบี่ยงเบนไปจากความเป็นจริง (Drost, 2011)

แบบสอบถามเป็นเครื่องมืออันหนึ่งที่ได้รับความนิยมอย่างแพร่หลายสำหรับใช้ในการ รวบรวมข้อมูลในการวิจัยด้านสังคมศาสตร์ โดยในกระบวนการพัฒนาและการตรวจสอบเครื่องมือเพื่อลดความคลาดเคลื่อนจากการวัด มี 2 ดัชนีที่แสดงถึงความมีคุณภาพของเครื่องมือ คือ ความเที่ยงตรงและความน่าเชื่อถือ (Kimberlin & Winterstein, 2008; Taherdoost, 2016)

การสร้างเครื่องมือวิจัยในการวิจัยเชิงปริมาณมีความสำคัญเป็นอย่างมาก เพราะความจริง (objective) คำถามการวิจัย (research question) และสมมติฐาน (hypothesis) มีความเชื่อมโยงกัน ข้อค้นพบและการสรุปผลขึ้นอยู่กับประเภทข้อมูลที่รวบรวมมาและมีความเกี่ยวข้องกับคำถามที่ได้จากผู้ตอบคำถาม หากเครื่องมือวิจัยมีคุณภาพและความเที่ยงตรง ผลลัพธ์และข้อค้นพบก็มีคุณภาพและความถูกต้องด้วย เหมือนกับคำกล่าวที่ได้ยินกันอย่างแพร่หลาย คือ เอาขยะใส่เข้าไปก็ได้ขยะออกมา (garbage in, garbage out) (Kumar, 2014: 188-189)

เครื่องมือที่ใช้ในการวิจัยด้านสังคมศาสตร์ไม่สามารถวัดได้ค่าความถูกต้อง (accurate) 100 เปอร์เซ็นต์ แต่สิ่งที่สำคัญ คือ ทำอย่างไรให้ผลการวิจัยได้รับการยอมรับว่ามีความถูกต้อง หรือในกระบวนการวิจัยเรียกว่า ความเที่ยงตรงที่เกิดจากความสัมพันธ์ระหว่างคำถามการวิจัยกับความจริง (objectives) ที่ทำการศึกษา โดยใช้วิธีการทางสถิติในการคำนวณความสัมพันธ์ระหว่างคำถามการวิจัยและตัวแปรที่เป็นผลลัพธ์ (outcome variables) เพื่อแสดงถึงความสัมพันธ์ระหว่างความจริงที่ทำการศึกษากับ

คำถาม นอกจากเครื่องมือต้องมีเที่ยงตรงแล้วยังต้องมีความสามารถในการวัดอย่างคงเส้นคงวา (consistent) ในการวัดแต่ละครั้ง หรือที่เรียกว่า ความน่าเชื่อถือ (reliability) เมื่อนำเครื่องมือไปใช้ภายใต้สถานการณ์ที่เหมือนกัน หรือประชากรที่เหมือนกัน ต้องได้ผลลัพธ์ที่เหมือนกัน โดยวิธีการในการตรวจสอบความเที่ยงตรงมี 2 วิธี คือ การทดสอบและทดสอบซ้ำ (test/retest) โดยการเครื่องมือ 1 ชุด นำไปทดสอบ 2 ครั้ง แต่ภายใต้สถานการณ์ที่เหมือนกัน และการทดสอบแบบขนาน (parallel forms of the same test) โดยใช้เครื่องมือ 2 ชุด นำไปทดสอบภายใต้สถานการณ์ที่เหมือนกันกับประชากรที่เหมือนกันหรือประชากร 2 กลุ่มที่มีลักษณะเหมือนกัน แม้ว่าเครื่องมือมีความเที่ยงตรงและมีความน่าเชื่อถือแล้ว ก่อนทำการรวบรวมข้อมูลก็ต้องนำเอาเครื่องมือวิจัยไปทำการทดสอบก่อน (pre-testing) อีกครั้ง เพื่อตรวจสอบปัญหาเกี่ยวกับความเข้าใจของแต่ละคำถามจากผู้ตอบคำถาม (responders) ในพื้นที่วิจัยจริงกับกลุ่มคนที่มีลักษณะที่เหมือนกับประชากรที่ทำการศึกษา (Kumar, 2014: 212-219)

ความคลาดเคลื่อนที่เกิดขึ้นไม่สามารถกำจัดออกไปได้จากการวิจัย แต่สามารถลดผลกระทบที่เกิดขึ้นได้ เช่น การสร้างแบบสอบถามให้มีคุณภาพและชัดเจน พยายามติดต่อและขอคำตอบใหม่ หากไม่ได้รับคำตอบหรือปฏิเสธไม่ให้คำตอบ อาจต้องใช้วิธีการให้รางวัลหรือสิ่งจูงใจ การอบรมผู้สัมภาษณ์ ให้มีทักษะที่ดี และตรวจสอบข้อมูลให้ถูกต้องและสมบูรณ์ (Groves cited by Scheaffer, Mendenhall III, & Ott, 2006: 18-25; Thompson, 2012: 5)

ทาย (Thye, 2000) อธิบายว่า การประมาณการความน่าเชื่อถือ (estimating reliability) มี 2 แนวคิด ดังนี้

1. ทฤษฎีการทดสอบแบบเก่า (Classical test theory: CT) เป็นวิธีการประมาณการความน่าเชื่อถือจากความแปรปรวนของคะแนนที่เป็นจริง (true score variance) และความแปรปรวนของความคลาดเคลื่อน (error variance) โดย ลี เจ. ครอนบาค (Cronbach 1947 cited by Thye, 2000) ได้เสนอตัวแบบปัจจัย (factor model) สำหรับใช้อธิบายความแตกต่างของสัมประสิทธิ์ความน่าเชื่อถือซึ่งเป็นความแปรปรวนจากการทดสอบโดยรวมที่มีสาเหตุมาจากความสัมพันธ์ของสัดส่วน (portion) จากปัจจัยต่างๆในการทดสอบ คือ ปัจจัยแบบทั่วไป (general factors) เป็นความแปรปรวนที่เกิดจากทุกข้อคำถาม ปัจจัย

แบบกลุ่ม (group factors) เป็นความแปรปรวนที่เกิดจากบางข้อคำถาม และปัจจัยแบบเฉพาะ (specific factors) เป็นความแปรปรวนที่เกิดจากข้อคำถามที่มีลักษณะพิเศษหรือไม่เหมือนข้อคำถามใดเลย

2. ทฤษฎีการสรุปอ้างอิง (Generalizability theory: GT) พัฒนาต่อยอดมาจากทฤษฎีการทดสอบแบบเก่าเพื่อใช้ในการประมาณการความน่าเชื่อถือที่เกี่ยวกับพฤติกรรม โดยมีฐานมาจากงานของ โฮเมอร์ ฮอยท์ (Hoyt, 1941 cited by Thye, 2000) ที่แสดงให้เห็นว่า ตัวแบบการวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) สามารถนำมาปรับใช้กับทดสอบการประมาณการความน่าเชื่อถือได้ด้วยวิธีการที่ไม่ยาก ไม่เหมือนกับทฤษฎีการทดสอบแบบเก่าที่ใช้ความแปรปรวนรวม (total variance) จากคะแนนที่เป็นจริง (true score) และความแปรปรวนของความคลาดเคลื่อน (error variance) เนื่องจากทฤษฎีการสรุปอ้างอิงแบ่งตัวความคลาดเคลื่อน (error term) ออกเป็นหลายแหล่งของความคลาดเคลื่อน โดยแยกความคลาดเคลื่อนที่เกิดจากความสัมพันธ์ระหว่างความแปรปรวนของปัจจัยเฉพาะกับข้อคำถามที่แตกต่างกัน เช่น แยกความคลาดเคลื่อนแบบเฉพาะ (specific error) ออกจากความคลาดเคลื่อนที่เป็นความสัมพันธ์จากการทดสอบเหตุการณ์ที่แตกต่างกัน เช่น ความคลาดเคลื่อนแบบชั่วคราว (transient error)

ทฤษฎีการสรุปอ้างอิงยังมีจุดแข็งในเรื่องของการใช้ตัวแบบเชิงสถิติสำหรับการสรุปอ้างอิง (generalized statistical models) ที่นิยมใช้กันมากในการวิจัยที่มีฐานมาจากการวิเคราะห์ความแปรปรวน (ANOWA) โดยการใช้วิธีการสกัด (extract) ผลรวมค่ากำลังสองของสารสนเทศ (sums of squares information) เพื่อตรวจสอบสัดส่วนของการประมาณการความแปรปรวนของแต่ละแหล่งข้อมูลที่ส่งผลให้เกิดความแปรปรวนรวมจาก 3 ส่วนประกอบ คือ ความแปรปรวนจากผู้ตอบแบบสอบถาม ความแปรปรวนจากข้อคำถาม และความแปรปรวนจากผู้ตอบแบบสอบถามกับข้อคำถาม

การรวบรวมข้อมูลเชิงปริมาณส่วนใหญ่ใช้แบบสอบถามแบบมีโครงสร้าง อาจให้ผู้ตอบแบบสอบถามเป็นผู้ตอบด้วยตนเอง (self-administered) หรือนักวิจัยเป็นผู้สัมภาษณ์ (interviewer-administered) หรืออาจเก็บรวบรวมข้อมูลจากการสังเกต (observations) และบันทึกลงในเครื่องมือวัดหรือแบบฟอร์ม ข้อมูลที่ได้มาสามารถแบ่งเป็น 2 ประเภท คือ ข้อมูลแบบค่าแบ่งกลุ่ม (categories) และ

ข้อมูลแบบค่าต่อเนื่อง (continuous) หรือแบ่งเป็น 4 มาตรา (scales) คือ นามมาตรา (nominal) อันดับมาตรา (ordinal) ช่วงมาตรา (interval) และอัตราส่วนมาตรา (ratio)

เป็นที่ทราบกันดีว่า ข้อมูลที่ใช้วิเคราะห์ด้วยสถิติแบบมีพารามิเตอร์ (parametric statistics) ต้องเป็นข้อมูลแบบเมตริก (metric data) ส่วนการวิเคราะห์ด้วยสถิติแบบไม่มีพารามิเตอร์ (nonparametric statistics) เป็นข้อมูลแบบไม่ใช่เมตริก (nonmetric data) หากละเมิด (violate) ข้อตกลง (assumptions) การค้นพบหรือข้อสรุปที่ได้จะมีความไม่ถูกต้อง (Verma, & Abdel-Salam, 2019: 3-4)

ประชากรและกลุ่มตัวอย่าง

ข้อมูลที่นำมาใช้ในการวิจัยด้านสังคมศาสตร์ส่วนใหญ่รวบรวมมาจากตัวอย่าง (samples) เพราะเป็นเรื่องยากมากที่นักวิจัยจะสามารถรวบรวมข้อมูลจากประชากรที่สนใจ (target populations) จำนวนทั้งหมด (entire/total/whole populations) มาทำการศึกษา เนื่องจากมีข้อจำกัดทั้งด้านเวลาและงบประมาณ ยกเว้นประชากรมีขนาดเล็ก (small populations) หรือการเลือกตัวอย่าง/สุ่มตัวอย่าง (sampling) อาจทำให้ความน่าเชื่อถือของผลการวิจัยลดลง เช่น การสำรวจความคิดเห็นทางการเมือง จึงต้องทำการรวบรวมข้อมูลจากประชากรทั้งหมด (census) เพื่อสร้างความน่าเชื่อถือ (reliability) และความเชื่อใจได้ (credibility) อย่างไรก็ตามเนื่องจากขนาดตัวอย่างมีความสัมพันธ์กับความคลาดเคลื่อน กล่าวคือ จำนวนขนาดตัวอย่างที่เพิ่มขึ้นจะทำให้ความคลาดเคลื่อนที่เป็นผลมาจากการเลือกตัวอย่าง (sampling error) จะลดลง แต่จำนวนขนาดตัวอย่างเพิ่มขึ้นจะทำให้เกิดความเคลื่อนที่ไม่ได้เป็นผลมาจากการเลือกตัวอย่าง (nonsampling error) เพิ่มขึ้นตามไปด้วย ดังนั้นขนาดของตัวอย่างจึงต้องมีจำนวนที่เพียงพอและสัมพันธ์กับเป้าหมายของการศึกษา และต้องมีขนาดใหญ่พอ (big enough) และมีผล (effect) ทำให้เกิดขนาด (magnitude) ของนัยสำคัญเชิงวิทยาศาสตร์หรือที่เรียกว่า นัยสำคัญทางสถิติ (statistically significant) แต่หากมีขนาดไม่ใหญ่พอและมีผลเชิงวิทยาศาสตร์เล็กน้อยก็ต้องสามารถตรวจสอบ (detectable) ด้วยวิธีการเชิงสถิติได้ (Yamane, 1967: 6; Cochran, 1977: 1-2; Henry, 1990: 12; Lenth, 2001: 188; Arnab, 2017: 469)

การเลือกตัวอย่างมีวัตถุประสงค์เพื่อใช้ในการอ้างอิง (inference) จากตัวอย่างไปสู่ประชากรที่ทราบจำนวน (finite population) ด้วยวิธีการประมาณ (estimate) จากค่าพารามิเตอร์ของประชากร

(population parameters) เช่น จำนวนรวม (total) ค่าเฉลี่ย (mean) และสัดส่วน (proportion) หากกระบวนการการออกแบบ (design) และการดำเนินการในการเลือกตัวอย่างไม่ถูกต้อง นอกจากจะมีผลต่อความเที่ยงตรงภายนอก (external validity) หรือความสามารถในการสร้างข้อสรุป (generalizability) ไปสู่ประชากรเป้าหมายแล้ว ยังมีผลต่อการนำเอาผลการศึกษาจากการค้นพบ (study finding) ไปใช้สร้างความสัมพันธ์เชิงสาเหตุ (causal relationship) ในสถานการณ์อื่นๆ หรือประชาชนกลุ่มอื่นๆ รวมถึงมีผลกระทบโดยตรงต่อความถูกต้องในการสรุปเชิงสถิติอีกด้วย (Scheaffer, Mendenhall III, & Ott, 2006: 7; Thompson, 2012: 2-6)

การทดสอบเชิงสถิติโดยทั่วไปเป็นการตรวจสอบความสัมพันธ์จากสิ่งที่ทำการสังเกตหรือสิ่งที่เกิดการเปลี่ยนแปลง ซึ่งการทดสอบนั้นมีความไวทั้งขนาดอิทธิพล (effect size) และขนาดของตัวอย่าง (sample size) ทำให้ได้ข้อสรุปที่เป็นเท็จ (false) ในการตัดสินใจปฏิเสธสมมติฐานหลัก (null hypothesis: H_0) ดังนั้นความแม่นยำ (precision) ของตัวประมาณการ (estimator) หรือความคลาดเคลื่อนในการเลือกตัวอย่างจึงมีความเกี่ยวข้องกับตัวอย่างทั้งด้านของขนาดตัวอย่างและวิธีการเลือกตัวอย่าง

ขนาดของตัวอย่างเป็นเครื่องมืออันหนึ่งในการสร้างความน่าเชื่อถือในการวิจัย แต่ก็มีความเกี่ยวข้องกับหลายปัจจัย กล่าวคือ กลุ่มตัวอย่างที่มีขนาดเล็ก (small sample size) อาจทำให้เกิดการไม่ปฏิเสธความจริงที่เป็นเท็จหรือที่เรียกว่า ความคลาดเคลื่อนแบบที่ 2 เพราะขนาดตัวอย่างมีจำนวนไม่เพียงพอที่จะทำให้เกิดผลกระทบหรือเกิดความแปรปรวนร่วมอย่างมีนัยสำคัญ และขณะเดียวกันอาจขาดความน่าเชื่อถือในบางบริบททางสังคม เช่น ความคิดเห็นทางการเมือง เป็นต้น การเพิ่มจำนวนตัวอย่างก็ทำให้เสียเวลาและค่าใช้จ่าย แม้ว่าความเที่ยงตรงและความน่าเชื่อถือในการวัด (reliability of measures) จะมีความสัมพันธ์กับขนาดกลุ่มตัวอย่างที่มีขนาดใหญ่กว่า เนื่องจากจะช่วยแก้ไขปัญหาความแปรปรวน (variation) ที่เพิ่มขึ้นได้หรือทำให้ความแปรปรวนในการสุ่มลดลง แม้ว่ากลุ่มตัวอย่างมีขนาดใหญ่ แต่หากเครื่องมือวัดที่นำมาใช้ในการเก็บรวบรวมข้อมูลมีความน่าเชื่อถือน้อย หรือมีความลำเอียงในการเลือกตัวอย่าง (sampling bias) ข้อมูลที่ได้จากการสังเกตก็จะมีความแปรปรวนมากขึ้น เมื่อข้อมูลมีความแปรปรวนมากขึ้นก็จะมีผลทำให้เกิดการปฏิเสธสมมติฐานหลักยากมากขึ้นตามมาแม้ว่าในความเป็นจริงจะมี

ความสัมพันธ์เกิดขึ้นจริง ดังนั้นในการออกแบบเลือกตัวอย่างจึงต้องเลือกใช้วิธีการที่เหมาะสม และได้คำตอบจากตัวแทนของประชากรอย่างแท้จริง (Henry, 1990: 9-59; Blair & Blair, 2015: 9-11)

ด้วยเหตุที่จำนวนตัวอย่างที่มีขนาดใหญ่ขึ้นหรือมีจำนวนเพิ่มขึ้นไม่สามารถขจัดความลำเอียงอย่างเป็นระบบ (systematic bias) ออกไปได้ จึงต้องใช้วิธีการเลือกตัวอย่างหรือการออกแบบเลือกตัวอย่าง (sampling design) เป็นกระบวนการที่ทำให้ได้มาแต่ละตัวอย่าง (*s*) ด้วยความน่าจะเป็น P(*s*) การเลือกตัวอย่างแบบสุ่ม (randomly) จะช่วยป้องกันไม่ให้เกิดความลำเอียง (bias) ส่วนตัวอย่างที่มาจากการสุ่มแบบตามสะดวก (convenience sampling) จะมีความลำเอียงสูงมาก และเพื่อให้มั่นใจว่ามีความสุ่ม (randomness) ควรใช้ตารางเลขสุ่ม (random number tables) เลือกตัวอย่างจากชุดรายชื่อ (set of names) ของประชากรเป้าหมาย แต่ในความเป็นจริงอาจมีเพียงการอธิบายขั้นตอนในการเลือกหน่วยตัวอย่างมากกว่าวิธีการเลือกที่ได้มาด้วยความน่าจะเป็นจากประชากรทั้งหมด ดังนั้น การลดความคลาดเคลื่อนในการประมาณจึงต้องมีความรอบครอบในการออกแบบเลือกตัวอย่างและใช้วิธีการประมาณที่เหมาะสมจากลักษณะของประชาการ เช่น ค่าเฉลี่ยของประชากร (population mean) หรือจำนวนรวมของประชาการ (population total) โดยไม่ใช้ข้อสันนิฐาน (assumption) เกี่ยวกับประชากรด้วยตัวนักวิจัยเอง (Flynn, Sakakibara, Schroeder, Bates, & Flynn, 1990: 260; Thompson, 2012: 2; Blair & Blair, 2015: 9-11)

การให้ได้มาซึ่งสิ่งที่ต้องการศึกษาหรือหน่วย (elements) ของประชากรมีวิธีการที่แตกต่างกันตามความเชื่อเกี่ยวกับการกระจายของตัวอย่าง ดังนี้ (Yamane, 1967: 3-6; Cochran, 1977: 8-9)

1. การเลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น (Nonprobability sampling) หรือเรียกว่าการเลือกตัวอย่างแบบดั้งเดิม (classical sampling theory) โดยใช้วิธีการเลือกตัวอย่างแบบเจาะจง (judgment/purposive sampling) การเลือกตัวอย่างแบบโควตา (quota sampling) และการส่งแบบสอบถามทางไปรษณีย์ (mail questionnaire)

2. การเลือกตัวอย่างแบบใช้หลักความน่าจะเป็น (Probability sampling) หรือเรียกว่า ทฤษฎีการสำรวจตัวอย่าง (sample survey theory) ได้แก่ การเลือกตัวอย่างแบบสุ่มอย่างง่าย (simple random sampling) การเลือกตัวอย่างแบบสุ่มด้วยการแบ่งชั้น/ชั้นภูมิ (stratified random sampling) การเลือก

ตัวอย่างแบบสุ่มอย่างเป็นระบบ (systematic random sampling) และการเลือกตัวอย่างแบบสุ่มเชิงพื้นที่ (cluster random sampling)

ขนาดตัวอย่าง (sample size) เป็นอีกหนึ่งองค์ประกอบที่สำคัญในการออกแบบการวิจัย เพราะมีผลอย่างมีนัยสำคัญต่อความเที่ยงตรง ความเสี่ยงของการเกิดความคลาดเคลื่อน (risk of errors) และข้อค้นพบ (finding) ที่ได้จากการวิจัย (Burmeister & Aitken, 2012) ปัจจัยที่มีความสัมพันธ์กับการคำนวณขนาดตัวอย่างมีดังนี้ (Youssef, 2011)

1. ขนาดของอิทธิพล (Effect size) คือ ค่าทางสถิติที่ได้จากทดสอบเปรียบเทียบหรือทดสอบความสัมพันธ์ เช่น ค่าเฉลี่ย ค่าสัมประสิทธิ์สหสัมพันธ์ ค่าสัมประสิทธิ์การถดถอย หากต้องการลดขนาดของอิทธิพล ต้องเพิ่มขนาดกลุ่มตัวอย่าง (effect size is decreased, sample size increases)

2. ความแปรปรวน (Variance) คือ ค่าความเบี่ยงเบน (standard deviation) ของความแปรปรวนที่ได้จากการวัดหรือตัวแปรตาม หากต้องการลดความแปรปรวน ต้องลดขนาดกลุ่มตัวอย่าง (variance is increased, sample size increases)

3. อำนาจจำแนกของสถิติ (Statistical power) คือ ดัชนีที่แสดงถึงจำนวนหรือร้อยละหรือสัดส่วน (fraction) ที่ทำให้เกิดนัยสำคัญทางสถิติ (statistically significant effect) แสดงถึงความน่าเชื่อถือสูงและมีความความคลาดเคลื่อนหรือโอกาสผิดพลาด (α) ในการสรุปผลที่ต่ำ หากต้องการเพิ่มอำนาจจำแนกของสถิติ (ลด β-error) ต้องเพิ่มขนาดกลุ่มตัวอย่าง (power is increased, sample size increases)

4. ระดับนัยสำคัญ (Level of significance) คือ ค่าความน่าจะเป็น (p-value) หรือระดับนัยสำคัญที่กำหนดเป็นเกณฑ์ในการยอมรับหรือปฏิเสธสมมติฐาน หากต้องการลดค่าความน่าจะเป็น ต้องเพิ่มขนาดกลุ่มตัวอย่าง (P value is decreased, sample size increases)

5. การวิเคราะห์ด้วยสถิติแบบหางเดียวหรือสองหาง (One or two-tailed statistical analysis) คือ การวิเคราะห์แบบหางเดียวประชากรกลุ่มหนึ่งต้องมีขนาดใหญ่กว่าอีกกลุ่มหนึ่ง ส่วนการวิเคราะห์แบบสองหางประชากรกลุ่มหนึ่งอาจมีขนาดมากกว่าหรือน้อยกว่าอีกกลุ่มหนึ่งก็ได้

เนื่องจากค่านัยสำคัญ (α) กับค่าอำนาจจำแนก (β) มีความสัมพันธ์เชิงผกผันกัน กล่าวคือ หากค่านัยสำคัญมีขนาดเล็กหรือมีความเสี่ยงน้อยที่จะปฏิเสธสมมติฐานหลักที่เป็นจริง (type I error) แต่ทำให้ค่าอำนาจจำแนกมีขนาดใหญ่หรือมีอำนาจจำแนกต่ำหรือมีความเสี่ยงมากที่จะยอมรับสมมติฐานหลักที่เป็นเท็จ (type II error) และถ้าค่าอำนาจจำแนกมีขนาดเล็กหรือมีอำนาจจำแนกสูงหรือมีความเสี่ยงน้อยที่จะยอมรับสมมติฐานหลักที่เป็นเท็จ แต่ค่านัยสำคัญจะมีขนาดใหญ่หรือมีความเสี่ยงมากที่จะปฏิเสธสมมติฐานหลักที่เป็นจริง จึงเกิดภาวะกลืนไม่เข้าคายไม่ออก เพราะหากต้องการป้องกันการเกิดความคลาดเคลื่อนแบบที่ 1 จึงต้องเลือกกำหนดค่า α ให้ต่ำ เช่น 0.01 หรือ 0.001 ซึ่งจะมีผลทำให้ต้องใช้ขนาดตัวอย่างเพิ่มขึ้น แต่ขณะเดียวกันจะทำให้เพิ่มโอกาสการเกิดความคลาดเคลื่อนแบบที่ 2 มากขึ้น และหากต้องการป้องกันการเกิดความคลาดเคลื่อนแบบที่ 2 ต้องเลือกกำหนดค่า β ให้ต่ำ เช่น 0.20 หรือ .10 ซึ่งมีผลทำให้ต้องใช้ขนาดตัวอย่างเพิ่มขึ้นเช่นเดียวกัน ดังนั้นนักวิจัยส่วนใหญ่จึงนิยมกำหนดค่า α=0.05 และค่า β=0.20 (power=0.80) เพราะจะทำให้ใช้จำนวนตัวอย่างไม่มากและไม่น้อยเกินไป (Cohen, 1977: p. 1-17)

ความแม่นยำของการประมาณค่า (precision of estimate) ถ้าต้องการความแม่นยำมาก ต้องใช้ตัวอย่างขนาดใหญ่ แต่ทฤษฎีหลักแนวโน้มเข้าสู่ศูนย์กลาง (Central Limit Theorem: CLT) เชื่อว่า ขนาดตัวอย่างตั้งแต่ 30 ตัวอย่างขึ้นไป น่าจะทำให้ข้อมูลมีการแจกแจงปกติ และโดยทั่วไปการกำหนดขนาดตัวอย่างจะพิจารณาจากขั้นของการเลือกใช้สถิติในการวิเคราะห์ผลการวิจัย เพราะหากขนาดตัวอย่างไม่เหมาะสมจะทำให้การอ้างอิง/อนุมานจากตัวอย่างไม่น่าเชื่อถือและอาจทำให้เกิดการสรุปผลที่ผิด

โดยทั่วไปสรุปกันว่า การกำหนดขนาดตัวอย่างมีปัจจัยหลักที่สำคัญ 3 ปัจจัย คือ ระดับของความแม่นยำ (level of precisions) ระดับของความเชื่อมั่นหรือความเสี่ยง (level of confidence or risk) และระดับของความแปรปรวน (degree of variability) ที่เกิดจากการวัด (Israel, 1992) แต่การวิจัยพื้นฐาน (basic research) ที่มีวัตถุประสงค์หลัก คือ การทดสอบสมมติฐาน (hypothesis test) ที่เน้นการยืนยัน (confirmatory) หรือตรวจสอบ (verify) ทฤษฎี นักวิจัยอาจต้องตัดสินใจเลือกวิธีคำนวณขนาดตัวอย่าง (sample size calculate) ด้วยวิธีการใดวิธีการหนึ่ง (trade-offs) ระหว่างการทดสอบนัยสำคัญทางสถิติและทดสอบอำนาจจำแนกทางสถิติ (Goldstein, 1989: 253)

การกำหนดขนาดตัวอย่างเพื่อเป็นตัวแทนจากประชากรเป้าหมายมีหลากหลายวิธี ดังนี้ (Israel, 1992; Bartlett, Kotrlik, & Higgins, 2001; Singh & Masuku, 2014)

1. การกำหนดขนาดตัวอย่างด้วยวิธีการทางสถิติ (Statistically determined sample size) แบ่งได้เป็น 2 วิธี คือ วิธีการกำหนดช่วงความเชื่อมั่น (confidence interval) และวิธีการกำหนดขนาดผลกระทบ (effect size) โดยมีทั้งแบบตารางสำเร็จรูป สูตรคำนวณ และใช้โปรแกรมคอมพิวเตอร์ ดังนี้

1.1 การใช้ตาราง (Tables) เป็นการกำหนดขนาดตัวอย่างโดยเลือกใช้ตารางสำเร็จรูปที่มีเงื่อนไขการคำนวณที่ตรงกับความต้องการ เช่น ระดับของความถูกต้อง ระดับของความน่าเชื่อถือ และระดับของความแปรปรวน เช่น ตารางขนาดตัวอย่างของทาโร่ ยามาเน่ (Taro Yamane) (1967: 398-399) ตารางขนาดตัวอย่างของโรเบิร์ต วี เครจซี่ (Robert V. Krejcie) และดาเรล ดับเบิลยู มอร์แกน (Daryle W. Morgan) (1970: 608) ตารางขนาดตัวอย่างของเจมส์ อี บาร์ตเลต (James E. Bartlett) โจ ดับเบิลยู โคทลิค (Joe W. Kotrlik) และแชดวิก ซี ฮิกกิ้น (Chadwick C. Higgins) (Bartlett, Kotrlik, & Higgins, 2001: 48)

1.2 การใช้สูตร (Formulas) มีสูตรสำหรับคำนวณขนาดตัวอย่างหลายวิธี อาจแบ่งได้เป็น 2 วิธี คือ วิธีที่เน้นความแม่นยำ (precision-based) ด้วยการประมาณการ (estimate) ความแตกต่างระหว่างค่าสัดส่วน (proportion) หรือค่าเฉลี่ย (mean) และวิธีที่เน้นอำนาจจำแนก (power-based) เพื่อตรวจสอบ (detect) หรือวัดระดับความมั่นใจ (degree of certainty) โดยมี 2 เงื่อนไขในการเลือกใช้ คือ ประเภทการวัด ข้อมูลที่ได้จากตัวแปรวัดเป็นข้อมูลแบ่งกลุ่ม (categorical data) หรือข้อมูลต่อเนื่อง (continuous/quantitative data) และขนาดของประชากร ประชากรมีขนาดใหญ่หรือไม่ทราบขนาดประชากร และประชากรมีขนาดเล็กหรือทราบขนาดประชากร เช่น สูตรของทาโร่ ยามาเน่ (Taro Yamane) (1967: 100) สูตรของโรเบิร์ต วี เครจซี่ (Robert V. Krejcie) และดาเรล ดับเบิลยู มอร์แกน (Daryle W. Morgan) (1970: 607) และสูตรของวิล เลี่ยม จี โครแครน (William G. Cochran) (1977: 75-80)

1.3 การใช้โปรแกรม (Software) ปัจจุบันมีทั้งโปรแกรมคอมพิวเตอร์ (software packages) และเว็บไซต์ (websites) สำหรับใช้คำนวณขนาดตัวอย่างให้เลือกใช้จำนวนมาก เช่น GPOWER, PASS,

Power and Precision, nQuery Advisor, Stat Power, Statpages website (http://statpages.org/#Power), Raosoft website (http://www.raosoft.com/samplesize.html) (Thomas & Krebs, 1997: 5; McCrum-Gardner, 2010: 12-13)

2. การกำหนดขนาดตัวอย่างด้วยวิธีการไม่ใช้หลักสถิติ (Nonstatistically determined sample size) เป็นการกำหนดขนาดตัวอย่างโดยขึ้นอยู่กับการตัดสินใจของนักวิจัย เพราะมีหลายปัจจัยที่เกี่ยวข้อง เช่น เวลา งบประมาณ วิธีการวิจัย บริบททางสังคม มีวิธีการดังนี้

2.1 การสำมะโน (Census) เป็นวิธีการที่ใช้ประชากรทั้งหมดเป็นตัวอย่าง ทำให้หมดปัญหาเกี่ยวกับความคลาดเคลื่อนในการเลือกตัวอย่างเพราะได้ข้อมูลจากทุกหน่วยของประชากร นิยมใช้กับประชากรที่มีขนาดเล็กหรือไม่เกิน 200 หน่วย และต้องการระดับของความถูกต้อง แต่อาจไม่เหมาะกับกรณีที่มีประชากรขนาดใหญ่เพราะต้องใช้งบประมาณและเวลามาก

2.2 การเลียนแบบ (Imitating) เป็นวิธีการกำหนดขนาดตัวอย่างเท่ากับเรื่องที่เคยศึกษามาแล้ว อาจเป็นการศึกษาซ้ำเรื่อง ซ้ำช่วงเวลา หรือซ้ำสถานที่ หากไม่ได้ทำการทบทวนวิธีการที่ใช้ในเรื่องที่เคยศึกษามาก่อน อาจต้องเผชิญกับความเสี่ยงในการเกิดความคลาดเคลื่อนซ้ำ

2.3 การกำหนดแบบคร่าวๆ (Rule of thumb) เช่น 5% ของประชากร 10 ตัวอย่างต่อ 1 ตัวแปร

สถิติและการวิเคราะห์ข้อมูลเชิงปริมาณ

สถิติที่ใช้ในการวิจัยด้านสังคมศาสตร์ส่วนใหญ่เป็นตัวเชื่อมจากคำถามการวิจัยไปสู่ข้อสรุป (conclusion) การทดสอบ (testing) และการนำเสนอหรือตีความ (interpret) โดยนักวิจัยจะเริ่มจากคำถามที่ต้องการตอบหรือสมมติฐานที่ต้องการตรวจสอบ (verify) และเปลี่ยนคำถามหรือสมมติฐานให้อยู่ในรูปของข้อมูล (data) หรือเหตุการณ์ (evidences) ที่สามารถสังเกตหรือวัดได้ โดยสถิติจะเข้ามาทำหน้าที่ช่วยให้นักวิจัยทำความเข้าใจ (simplify) หรือสรุป (summarize) ข้อมูลที่เป็นตัวเลขที่มีจำนวนมากให้ได้สารสนเทศที่สั้น/กระชับและมีความแม่นยำทำให้การตรวจสอบทำได้ง่ายมากขึ้นและช่วยตัดสินใจได้อย่างเพียงพอ

โดยทั่วไปและตำราหลายเล่มแบ่งประเภทสถิติตามเทคนิคทางสถิติเป็น 3 ประเภท คือ สถิติเชิงพรรณนา สถิติเชิงอ้างอิง และสถิติเชิงความสัมพันธ์และทำนาย (Healey, 1999: 259-260; Field, 2009: 540; Yang, 2010: 7; Evans, 2014: 2-5) ดังนี้

1. สถิติเชิงพรรณนา (Descriptive statistics) คือ สถิติสาขาหนึ่งที่มีวัตถุประสงค์ในการใช้บรรยาย (describes) ลักษณะของข้อมูลพื้นฐานที่ใช้ในการศึกษา โดยใช้วิธีการสรุป (summary indictors) หรือลดจำนวนข้อมูลขนาดใหญ่จากกลุ่มตัวอย่างหรือการวัดให้อยู่ในรูปของดัชนีหรือค่าสรุปขนาดเล็กที่มีความกระชับหรือสามารถเข้าใจได้ ได้แก่ ค่าความถี่ (frequency) ค่าร้อยละ (percentage) ค่าฐานนิยม (mode) ค่ามัธยฐาน (median) ค่าเฉลี่ย (mean) พิสัย (range) ส่วนเบี่ยงเบนมาตรฐาน (standard deviation) ส่วนเบี่ยงเบนควอไทล์ (quartile deviation) ส่วนเบี่ยงเบนเฉลี่ย (mean or average deviation) ค่าแปรปรวน (variance) สัมประสิทธิ์ความแปรปรวน (coefficient of variation) ควอไทล์ (quartiles) เดไซล์ (decile) เปอร์เซ็นไทล์ (percentiles) N-ไทล์ (N-tiles) และคะแนนมาตรฐาน (standardized scores: z-scores)

2. สถิติเชิงอ้างอิง (Inferential statistics) คือ สถิติกลุ่มตัวแบบเชิงเส้นทั่วไป (general linear models) ที่คำนวณข้อมูลจากกลุ่มตัวอย่างหนึ่งกลุ่มตัวอย่าง หรือมากกว่าหนึ่งกลุ่มตัวอย่าง แล้วให้ค่าสถิติ (เช่น ค่าเฉลี่ยของประชากรทั้งหมด) ที่สามารถใช้อ้างอิงหรือทำนายลักษณะของประชากร เพื่อตัดสินใจเชิงความน่าจะเป็น (probability) โดยทั่วไปมีการแบ่งเป็น 2 ประเภท ดังนี้

2.1 สถิติเชิงอ้างอิงแบบมีพารามิเตอร์ (Parametric statistics) คือ สถิติที่มีค่าตัวเลขทางประชากรที่ใช้ในการอ้างอิงไปสู่คุณลักษณะ (characteristics) ของประชากร ข้อตกลงเบื้องต้นข้อมูลที่ใช้ในการวิเคราะห์มีระดับการวัดเป็นแบบค่าต่อเนื่อง ได้แก่ ช่วงมาตรา (interval) และสัดส่วนมาตรา (ratio) และข้อมูลมีการแจกแจงปกติ (normal distribution) เช่น การทดสอบค่าเฉลี่ยกลุ่มตัวอย่าง 1 กลุ่มด้วยค่าที (One-sample t-test) การทดสอบสอบค่าเฉลี่ยระหว่าง 2 กลุ่มตัวอย่างด้วยค่าที (Two-sample t-test) การทดสอบค่าเฉลี่ยระหว่างกลุ่มตัวอย่างที่มีความสัมพันธ์กันด้วยค่าที (Paired t-test) การวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA) การวิเคราะห์ความแปรปรวนแบบสองทาง (Two-way ANOVA) การวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measures ANOVA) การวิเคราะห์

สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) การวิเคราะห์การถดถอยเชิงเส้นตรง (Linear Regression) การวิเคราะห์องค์ประกอบ (Factor Analysis) และการวิเคราะห์ตัวแปรพหุ (Multivariate Analysis)[1]

2.2 สถิติเชิงอ้างอิงแบบไม่มีพารามิเตอร์ (Nonparametric statistics) คือ สถิติที่ไม่มีค่าตัวเลขทางประชากรที่ใช้ในการอ้างอิงไปสู่คุณลักษณะ (characteristics) ของประชากร จึงทำให้บางตำราจัดรวมอยู่ในกลุ่มของสถิติเชิงพรรณนา ข้อมูลที่ใช้ในการวิเคราะห์มีระดับการวัดเป็นแบบข้อมูลแบ่งกลุ่ม ได้แก่ นามมาตรา (nominal) และอันดับมาตรา (ordinal/rank-order) บางตำราอาจเรียกสถิติประเภทนี้ว่า การทดสอบที่มีการแจกแจงแบบอิสระ (distribution-free test)[2] หรือ การทดสอบที่เป็นอิสระจากข้อตกลง (assumption-free test) เพราะสถิติที่ใช้ทดสอบข้อมูลอาจมีข้อตกลงเพียงเล็กน้อยหรือไม่มีข้อตกลงเลยเกี่ยวกับลักษณะการแจกแจงของข้อมูล (data distributions) เช่น การทดสอบความสอดคล้องของข้อมูลด้วยไคสแควร์ (X^2 goodness of fit) การทดสอบความเป็นอิสระของข้อมูลด้วยไคสแควร์ (X^2 Test for Independence) การทดสอบค่ากลางของข้อมูล 1 กลุ่มตัวอย่าง (One-sample Test) การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน (2 Related Samples Test) การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน (Independence Sample Test) การทดสอบสหสัมพันธ์แบบสเปียร์แมน (Spearman Correlation Test) การวิเคราะห์การถดถอยแบบไม่มีพารามิเตอร์ (Nonparametric Regression Analysis) และการวิเคราะห์พหุแบบไม่มีพารามิเตอร์ (Nonparametric Multivariate Analysis)[3]

[1] การวิเคราะห์ที่มีเงื่อนไขต้องใช้ตัวอย่างขนาดใหญ่และข้อมูลมีการแจกแจงแบบปกติ เช่น ตัวแบบสมการโครงสร้างแบบใช้ค่าความแปรปรวนร่วม (Covariance Based Structural Equation Modeling: CB-SEM) ที่เน้นการวิเคราะห์เชิงยืนยัน (Awang, Afthanorhan, & Asri 2015)

[2] เป็นการทดสอบที่ข้อมูลมีลักษณะการแจกแจงแบบใดก็ได้ หรืออาจเรียกว่า การทดสอบที่ไม่มีข้อตกลงหรือการทดสอบที่เป็นอิสระจากข้อตกลง (free assumption)

[3] การวิเคราะห์ที่สามารถใช้ตัวอย่างขนาดเล็กและข้อมูลมีการแจกแจงแบบอิสระ เช่น การวิเคราะห์ตัวแบบสมการโครงสร้างแบบใช้ค่าความแปรปรวน (Variance Based Structural Equation Modeling: VB-SEM) ที่เน้นการวิเคราะห์เชิงสำรวจ (exploratory analysis) (Awang, Afthanorhan and Asri 2015)

3. สถิติเชิงความสัมพันธ์และการทำนาย (Correlational and predictive statistics) คือ สถิติที่ใช้ในการบอก/พรรณนา (describing) ความสัมพันธ์ระหว่างเหตุการณ์ต่างๆ หรือทำนายจากเหตุการณ์หนึ่งไปสู่อีกเหตุการณ์หนึ่ง โดยมีรายละเอียด ดังนี้

3.1 สถิติเชิงความสัมพันธ์ เป็นเครื่องมือที่ใช้ในการวัดความเข้มข้น (strength) ที่แตกต่างกันของความสัมพันธ์ (relationships) โดยรวมระหว่าง 2 ชุดหรือกลุ่มข้อมูลที่ได้จากการสังเกต เช่น ความสัมพันธ์ระหว่างความยากจนกับอาชญากรรม และนำไปใช้ในเชิงการอ้างอิงจากกลุ่มตัวอย่างไปสู่กลุ่มประชากร ได้แก่ การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) และการวิเคราะห์สหสัมพันธ์แบบสเปียร์แมน (Spearman Correlation)

3.2 สถิติเชิงการทำนาย เป็นเครื่องมือที่ใช้ในการทำนาย (prediction) ผลลัพธ์ของเหตุการณ์หนึ่งจากอีกเหตุการณ์หนึ่ง เช่น การทำนายความพร้อมด้านเทคโนโลยีสารสนเทศและการสื่อสารจากสถานะภาพทางเศรษฐกิจและสังคม การทำนายจะมีความถูกต้องมากน้อยเพียงใดขึ้นกับว่าข้อมูลที่ได้มาจากการสังเกตมีความสัมพันธ์กันมากน้อยเพียงใด ได้แก่ การวิเคราะห์การถดถอยเชิงเส้นตรง (linear regression) และการวิเคราะห์การถดถอยเชิงโลจิสติก (logistic regression)

จากที่กล่าวมาจะเห็นว่าสถิติที่ใช้ในการวิเคราะห์ข้อมูลแต่ละประเภทมีความคาบเกี่ยวกัน การจำแนกประเภทต้องบอกหลักเกณฑ์ให้ชัดเจน เพราะสถิติเชิงความสัมพันธ์และเชิงการทำนายยังสามารถจำแนกได้เป็น 2 ประเภท คือ สถิติแบบมีพารามิเตอร์และสถิติแบบไม่มีพารามิเตอร์ หรือสถิติเชิงพรรณาสามารถจัดเป็นสถิติแบบไม่มีพารามิเตอร์ได้เช่นกัน เนื่องจากเป็นสถิติที่ใช้พรรณนาเฉพาะกลุ่มตัวอย่างหรือประชากรเท่านั้น โดยไม่อ้างอิงจากกลุ่มตัวอย่างไปสู่ประชากร

การตรวจสอบข้อมูลตามข้อตกลงเบื้องต้น

เป้าหมายที่สำคัญของการใช้สถิติ คือ การนำเอาสารสนเทศ (information) จากกลุ่มตัวอย่างอ้างอิงไปสู่การอธิบายหรือสรุปประชากร (Scheaffer, Mendenhall III, & Ott, 2006: 48-50) การละเมิด (violation) ข้อตกลงบางครั้งอาจไม่มีผลต่อข้อสรุปของการวิจัยแต่อย่างไร แต่บางครั้งอาจมีผลต่อการแปลความหมายของผลการวิจัยเป็นอย่างมาก และไม่น่าที่จะทำให้ได้ผลการวิจัยที่ถูกต้องและได้ข้อสรุปที่

น่าเชื่อถือ (Ghasemi & Zahediasl, 2012: 486-489) เกิดความคลาดเคลื่อนแบบที่ 1 (type I (alpha) error) และความคลาดเคลื่อนแบบที่ 2 (type II (beta) error) หรือทำให้การประมาณการค่านัยสำคัญ (significances) หรือขนาดอิทธิพล (effect sizes) ต่ำกว่าหรือมากกว่า แต่เรากลับยัดเยียดความถูกต้องไว้ในผลลัพธ์ ข้อสรุป และการยืนยันผลของการวิจัย โดยไม่ได้ทำการทดสอบตามหลักสถิติว่า เป็นไปตามข้อตกลงเบื้องต้นหรือไม่ (Osborne & Waters, 2002)

อย่างไรก็ตามยังมีนักวิจัยจำนวนไม่น้อยที่เชื่อเหมือนกับ เดวิด การ์สัน (Garson, 2012: 8) ที่อธิบายไว้ว่า สถิติเชิงอ้างอิงแบบมีพารามิเตอร์ เช่น การทดสอบสหสัมพันธ์ (correlation) การทดสอบการถดถอย (regression) การทดสอบค่าซี (Z-test) การทดสอบค่าที (T-test) และการทดสอบค่าเอฟ (F-test) นอกจากมีข้อตกลงเบื้องต้น คือ ข้อมูลมีค่าแบบต่อเนื่อง (continuous data) หรือมีการวัดตั้งแต่ระดับช่วงมาตรา (interval level) ขึ้นไป ข้อมูลต้องมีแจกแจงปกติ (normal distribution) และหากเป็นการวิเคราะห์ความแปรปรวน (analysis of variance) เพื่อเปรียบเทียบตัวอย่าง 2 กลุ่ม หรือมากกว่า 2 กลุ่ม ข้อมูลต้องมีความเป็นแบบเดียวกันของความแปรปรวน (homogeneity of variance) ส่วนการทดสอบด้วยสถิติแบบไม่มีพารามิเตอร์ เช่น การทดสอบความเป็นอิสระต่อกันของสองตัวแปรหรือไคสแควร์ (Chi-square test) ไม่มีข้อตกลงเกี่ยวกับการแจกแจงปกติ รวมถึงวิธีการทดสอบแบบไม่มีพารามิเตอร์ด้วยการประมาณการค่าด้วยวิธีบูตสแตรป (bootstrapped estimates) ที่ไม่มีข้อตกลงเบื้องต้นเกี่ยวกับการแจกปกติเช่นกัน

การตรวจสอบข้อมูลก่อนการวิเคราะห์หรือประมวลผลข้อมูลมีความสำคัญไม่น้อยไปกว่าขั้นตอนอื่นๆ กล่าวคือ การกำหนดชื่อตัวแปรและความหมายมีส่วนช่วยให้การแปรผลทำได้ง่าย หากมีการกำหนดความหมายค่าตัวแปรผิดก็มีผลต่อการตีความผิด และหากไม่กำหนดให้ค่าตัวแปรบางตัวเป็นค่าสูญหาย โปรแกรมก็จะนำค่าไปรวมในการประมวลผลทำให้ได้ผลลัพธ์ไม่ถูกต้อง และการกำหนดระดับการวัดก็มีผลต่อการเลือกคำสั่งประมวลผลทางสถิติในบางรายการ เพราะโปรแกรมจะไม่แสดงรายการตัวแปรที่มีระดับการวัดที่ไม่ตรงกับเงื่อนไขทางสถิติที่เลือก รวมถึงการป้องกันไม่ให้นำเอาตัวแปรไปใช้ในการประมวลผลผิด เช่น ใช้ตัวแปรเพศในการคำนวณหาค่าเฉลี่ย เป็นต้น

การตรวจสอบคุณลักษณะของข้อมูล (exploratory data analysis) มีความสำคัญต่อการเลือกใช้สถิติเชิงอ้างอิงหรือสถิติเชิงอนุมาน (inferential statistics) ดังนี้

1. สถิติแบบมีพารามิเตอร์ (Parametric statistics) ข้อมูลต้องเป็นข้อมูลตัวเลขหรือต่อเนื่องหรือมีการวัดตั้งแต่ระดับช่วงมาตรา (interval scale) ขึ้นไป มีการแจกแจงข้อมูลเป็นโค้งปกติ (normally distributed data) มีความเป็นแบบเดียวกันของความแปรปรวน (homogeneity of variance) ข้อมูลอย่างน้อยมีการระดับการวัดช่วงมาตรา (interval data) ความเป็นอิสระของข้อมูล (independence) และ/หรือมีความเป็นเส้นตรงในการวิเคราะห์การถดถอยเชิงเส้นตรง (linear regression analysis)

2. สถิติแบบไม่มีพารามิเตอร์ (Nonparametric statistics) เป็นสถิติที่มีข้อตกลงเพียงเล็กน้อยหรือไม่มีข้อตกลง (assumption-free test) ข้อมูลมีการวัดระดับใดก็ได้ ไม่จำเป็นต้องมีการแจกแจงเป็นโค้งปกติ ข้อมูลแต่ละกลุ่มไม่จำเป็นต้องมีการกระจายเท่ากัน และ/หรือไม่มีความเป็นเส้นตรงในการวิเคราะห์การถดถอยไม่ใช่เชิงเส้นตรง (nonlinear regression analysis)

การบันทึกข้อมูลผิดและคุณลักษณะของข้อมูลที่ไม่ปกติมีผลทำให้การประมวลผลได้ผลลัพธ์และการนำไปใช้ในการตีความผิดตามมา การตรวจสอบข้อมูลก่อนการประมวลผลไม่เพียงแต่เป็นการค้นหาข้อมูลที่ไม่ตรงกับข้อมูลที่เก็บมาจากกลุ่มตัวอย่างเท่านั้น แต่ข้อมูลที่มีความถูกต้องตรงกับแบบสอบถามอาจเป็นข้อมูลที่มีลักษณะของข้อมูลที่ไม่เป็นไปตามเงื่อนไขของการวิเคราะห์ด้วยสถิติก็เป็นได้ เช่น ข้อมูลมีการแจกแจงเป็นโค้งไม่ปกติ (asymmetric or nonnormal distribution) ข้อมูลมีความแปรปรวนไม่เท่ากัน (difference variance) ลักษณะของข้อมูลแบบนี้อาจมีผลทำให้การอธิบาย การตัดสินใจ และการพยากรณ์ที่อ้างอิงไปถึงประชากรไม่ถูกต้อง

นักวิจัยส่วนใหญ่ไม่ให้ความสนใจข้อตกลงของสถิติแบบมีพารามิเตอร์ ดังนั้นผลการวิเคราะห์จึงไม่มีความถูกต้อง การตรวจสอบข้อมูลให้ตรงกับข้อตกลง (met assumptions) ของสถิติทดสอบแต่ละประเภทจึงมีความสำคัญทั้งด้านความถูกต้องตามข้อเท็จจริงของข้อมูล และเป็นพื้นฐานของการตัดสินใจเลือกใช้สถิติเชิงอ้างอิงให้ถูกต้องตามหลักสถิติ การใช้สถิติแบบมีพารามิเตอร์ นอกจากต้องใช้ข้อมูลที่มีการกระจาย (distributions) จากบัญชีรายชื่อขนาดใหญ่ (large catalogues) แล้ว ยังต้องทำการตรวจสอบ

ข้อมูลเพื่อให้ตรงกับข้อตกลงเบื้องต้น เพื่อให้เกิดความน่าเชื่อถือในการสรุปและอ้างอิงไปถึงประชากร ดังนี้ (Allison, 2002: V; Hair & Others, 2006: 49-50; Field, 2009: 132-133; Garson, 2012: 13-48; Verma, & Abdel-Salam, 2019: 66-82)

1. การตรวจสอบค่าผิดปกติ (Outlier) สิ่งแรกที่นักวิจัยต้องทำก่อนอื่น คือ การตรวจสอบความผิดปกติภายในชุดข้อมูล เช่น ค่าที่กรอกผิด (wrong entry) ค่าสุดโต่ง (extreme) ซึ่งเป็นค่าที่ผิดปกติไปจากสังเกตหรือการสำรวจจากกลุ่มตัวอย่าง เพราะค่าผิดปกติอาจทำให้ข้อมูลมีการแจกแจงไม่ปกติ หรือค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) เพิ่มขึ้นหรือลดลง เป็นต้น

2. การตรวจสอบค่าสูญหาย (Missing) ความไม่สมบูรณ์ของข้อมูลเกิดมาจากหลายสาเหตุ เช่น ไม่รู้ ไม่ตอบ ไม่ทราบ ตอบไม่ได้ ลืมถาม ลืมตอบ เป็นต้น แฟ้มข้อมูลที่มีค่าสูญหายจำนวนมากอาจมีจำนวนข้อมูลไม่ถึงครึ่งของข้อมูลที่เก็บรวบรวมมา ซึ่งมีผลทำให้ได้คำตอบที่ไม่ได้มาจากตัวแทนของประชากรอย่างแท้จริง ค่าสถิติที่ได้ไม่สามารถนำไปใช้ในการทำนายหรืออธิบายได้อย่างถูกต้อง และเกิดการตัดสินใจผิดในการปฏิเสธหรือยอมรับสมมติฐาน

3. การตรวจสอบการแจกแจงปกติ (Normality) เป็นข้อสันนิฐานที่สำคัญเกี่ยวกับความน่าเชื่อถือของข้อมูลที่ใช้ในการสรุปและอ้างอิงไปถึงประชากร ถ้าข้อมูลมีการแจกแจงปกติมีโอกาสน้อยที่จะเกิดความลำเอียง แต่หากข้อมูลมีการแจกแจงไม่ปกติการทดสอบสมมติฐานและการสรุปก็จะไม่ถูกต้อง

4. การตรวจสอบความเป็นสุ่ม (Randomness) ข้อมูลที่ใช้ในการวิเคราะห์ไม่เพียงแต่จะได้มาจากตัวอย่างที่สุ่ม (random samples) เท่านั้น แต่ข้อมูลยังต้องมีลักษณะการจัดเรียงแบบสุ่ม (random sequences) โดยใช้การทดสอบความเป็นสุ่มของข้อมูลแบบรันส์ (runs test)

5. การตรวจสอบความเป็นแบบเดียวของความแปรปรวน (Homogeneity of variance) การทดสอบเปรียบเทียบค่าเฉลี่ยของ 2 กลุ่มตัวอย่าง ข้อมูลระหว่างกลุ่มตัวอย่างต้องมีความแปรปรวนเหมือนกัน

6. การตรวจสอบข้อมูลของตัวแปร ข้อมูลต้องมีการวัดอย่างน้อยระดับช่วงมาตรา (interval data) เป็นตัวแปรที่มีค่าแบบต่อเนื่อง (continuous data) เช่น ตัวแปรแบบช่วงมาตรา และตัวแปรแบบสัดส่วนมาตรา

7. การตรวจสอบความเป็นอิสระของข้อมูล (Independence) หรือ การตรวจสอบความสัมพันธ์ภายในตัวแปร (autocorrelation) ข้อมูลแข้อมูลจากแต่ละตัวอย่างต้องเป็นอิสระจากกันและไม่มีอิทธิพลซึ่งกันและกัน เช่น การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ (multiple linear regression) การวิเคราะห์การถดถอยเชิงโลจิสติก (logistic regression) และการวิเคราะห์การถดถอยเชิงปัวซอง (poisson regression)

8. การตรวจสอบความเป็นเส้นตรง (Linearity) ข้อมูลที่นำมาใช้ในการวิเคราะห์ คือ ตัวแปรต้องมีความสัมพันธ์ระหว่างกันเชิงเส้นตรง เช่น การวิเคราะห์การถดถอยเชิงเส้นตรง (linear regression) การวิเคราะห์ปัจจัย (factor analysis) และการวิเคราะห์ตัวแบบสมการเชิงโครงสร้าง (structural equation model)

การทดสอบสมมติฐาน

การทดสอบนัยสำคัญทางสถิติจากลักษณะของการแจกแจงของข้อมูล (data distributions) ที่นิยมใช้กันอย่างแพร่หลายในปัจจุบันมี 2 แนวทาง ดังนี้

1. การทดสอบนัยสำคัญทางสถิติ (Statistical significance test) มีฐานความคิดมาจาก โรนัลด์ ไอล์เมอร์ ฟิชเชอร์ (Ronald Aylmer Fisher) ที่ใช้ค่า *P* หรือความคลาดเคลื่อนที่น่าจะเป็น (probable errors) ที่ได้มาจากการวัดความแตกต่างกันระหว่างสองสิ่งที่ควรจะเหมือนกัน (discrepancy) ซึ่งมี 2 วิธี คือ การวัดที่ได้จากค่าที่เป็นจริงจากการสังเกต (observation) กับสมมติฐาน (hypothesis) และการวัดจากค่าที่ได้รับมาหรือนำมาใช้ (value adopted) กับค่าที่ได้ด้วยวิธีการของความควรจะเป็นสูงสุด (maximum likelihood) หากมีความกลมกลืนหรือแนบกันสนิท (goodness of fit) ระหว่างกันมาก ค่า *P* จะมีค่ามาก แต่หากค่า *P*=0.05 (เป็นขีดสูงสุดของค่าเบี่ยงเบนอย่างมีนัยสำคัญ) แสดงว่า เกิดการเบี่ยงเบนไปจากความคาดหวังอย่างมีนัยสำคัญที่ชัดเจน (Fisher, 1934: 21-23, 45, 83-85, 305-306)

การทดสอบนัยสำคัญทางสถิตินิยมใช้ในการทดสอบสมมติฐานเชิงอ้างอิงหรือการอนุมาน (inferential test) เพื่อการยืนยัน (confirmatory) มากกว่าการสำรวจ (exploratory) โดยใช้ค่า *P* ตัดสินใจปฏิเสธสมมติฐานหลัก (null hypothesis: H_0) ซึ่งเป็นประเด็นที่ทำให้เกิดความกังวลว่า การ

ปฏิเสธ (rejection) สมมติฐานหลัก (H_0) ถูกต้องมากน้อยเพียงใด จนเกิดคำถามหลายประเด็นตามมา เช่น การปฏิเสธสมมติฐานหลัก (H_0) มีความน่าเชื่อถือเพียงใด? การทดสอบทางสถิติมีอำนาจจำแนก (power) เท่าไร? การคาดการ (expectation) ว่า สมมติฐานหลัก (H_0) ที่เป็นเท็จนั้นเป็นจริง เกิดความคลาดเคลื่อนแบบที่ 2 (type II error) หรือไม่? (Cohen, 1962: 145)

2. การทดสอบอำนาจจำแนกทางสถิติ (Statistical power test) เป็นวิธีการหนึ่งทางสถิติในการทดสอบความน่าจะเป็น (probabilities) ในการปฏิเสธสมมติฐานหลัก (null hypothesis: H_0) ความน่าจะเป็น คือ ผลลัพธ์ที่ได้มาจากการสรุปว่า ปรากฏการณ์นั้นมีอยู่จริง อำนาจจำแนกของการทดสอบทางสถิติอิงอยู่กับ 3 พารามิเตอร์ (parameters) ดังนี้ (Cohen, 1977: 1-17)

2.1 เกณฑ์การกำหนดนัยสำคัญ (Significance criterion) เป็นพื้นที่วิกฤติของการปฏิเสธสมมติฐานหลัก (null hypothesis: H_0) และเป็นอัตราการเปรียบเทียบ (standard) การพิสูจน์ (proof) ความมีอยู่จริงของปรากฏการณ์ (phenomenon exist) ซึ่งอาจเรียกว่า ระดับนัยสำคัญ (significance level) ค่าอัลฟา (alpha value: α) ความคลาดเคลื่อนแบบที่ 1 (type I error) หรือ ความคลาดเคลื่อนแอลฟา (alpha error) เป็นต้น

2.2 ความน่าเชื่อถือที่ได้จากตัวอย่าง (Reliability of the sample results) คือ ความใกล้ (closeness) ที่ได้จากการคาดหมาย (expected) ด้วยการประมาณการ (approximate) ความเกี่ยวเนื่องกัน (relevant) ระหว่างค่าของตัวอย่าง (sample value) กับค่าของประชากร (population value) ความน่าจะเชื่อถืออาจเกี่ยวข้องหรืออาจไม่เกี่ยวข้องโดยตรงกับหน่วยของการวัด (unite of measurement) ค่าของประชากร และรูปแบบการกระจายของประชากร (shape of the population distribution) แต่มีความเกี่ยวข้องโดยตรงกับขนาดตัวอย่าง กล่าวคือ ขนาดตัวอย่างขนาดใหญ่จะทำให้ความคลาดเคลื่อนน้อยลง และมีความน่าเชื่อถือหรือความแม่นยำ (precision) มากขึ้น จึงกล่าวสรุปได้ว่ามีความสัมพันธ์ระหว่างขนาดตัวอย่างและอำนาจจำแนก เมื่อขนาดตัวอย่างใหญ่มากขึ้นจะทำให้อำนาจจำแนกทางสถิติเพิ่มขึ้น อย่างไรก็ตาม นักวิจัยต้องออกแบบและควบคุมปัจจัยที่เข้าไปมีผลกระทบกับความแปรปรวนของข้อมูลที่ได้จากตัวอย่าง เพื่อเพิ่มอำนาจจำแนกให้มากขึ้น

2.3 ขนาดอิทธิพล (Effect size) คือ ระดับความเข้ม (degree) ของปรากฏการณ์ (phenomenon) ที่เกิดขึ้นในประชากร หรือระดับความเข้มของสมมติฐานหลัก (H_0) ที่เป็นเท็จ เป็นค่าเฉพาะค่าหนึ่งที่ไม่ใช่ 0 (some specific nonzero values) เมื่อค่าขนาดอิทธิพลที่มีขนาดใหญ่กว่า แสดง

ว่า ระดับความเข้มของปรากฏการณ์ที่อยู่ภายใต้การศึกษาปรากฏออกมาให้เห็นชัดเจนมากกว่าค่าที่มีขนาดเล็กกว่า หรืออาจกล่าวได้ว่า เป็นร้อยละของการไม่ทับกัน (nonoverlap) ของคะแนนระหว่างกลุ่มที่ทดสอบ โดยปกติค่าขนาดอิทธิพลของสมมติฐานหลัก (H_0) ในการเปรียบเทียบสัดส่วนของประชากร การทดสอบค่าเฉลี่ยของตัวอย่าง และการเปรียบเทียบสหสัมพันธ์ (r) คือ 0

ความคลาดเคลื่อนของข้อมูลสามารถเกิดขึ้นได้เสมอ สาเหตุที่ทำให้เกิดความคลาดเคลื่อนมีความสัมพันธ์กัน อาจมาจากการเก็บข้อมูลในช่วงเวลาที่แตกต่างกัน และรวมถึงการเก็บข้อมูลในช่วงเวลาเดียวกันจากกลุ่มตัวอย่างที่มีความแตกต่างกัน ก็ทำให้เกิดความคลาดเคลื่อนมีความสัมพันธ์กันได้

องค์การอาหารและยา (Food and Drug Administration: FDA) ของประเทศสหรัฐอเมริกา เป็นหน่วยงานหนึ่งที่ให้ความสำคัญกับผลการศึกษาที่มีความไม่แน่นอน (uncertaintiy) หรือเป็นข้อมูลที่มีความบกพร่อง (imperfect data) จึงสนับสนุนให้มีการศึกษาและพัฒนาโปรแกรมในทดสอบการวิเคราะห์ความลำเอียงเชิงปริมาณ (quantitative bias analysis) ที่เกิดจากความคลาดเคลื่อนอย่างเป็นระบบ (systematic error) จากการใช้วิธีการที่คลาดเคลื่อนในการประมาณการ (estimate) การกำหนดขนาด (magnitude) และการวัด (measure) โดยรวบรวมข้อมูลหรือสารสนเทศที่จำเป็น เช่น ประเภทการออกแบบการศึกษา (study design type) ประเภทของความลำเอียงและพารามิเตอร์ (bias types and parameters) และประเภทของข้อมูล (data types) เพื่อนำไปใช้ในการอ้างอิงการตัดสินใจ (decision making) ด้วยความชัดเจนและโปร่งใสที่อยู่บนฐานที่เป็นวิทยาศาสตร์ (scientific basis) ในการดำเนินการด้านกฎหมายเกี่ยวกับอาหารและยา (Lash & Others, 2016)

ความคลาดเคลื่อนในการวิจัย

ความคลาดเคลื่อนโดยรวม (total error) มีสาเหตุมาจากความคลาดเคลื่อนที่ไม่ได้เกิดจากการเลือกตัวอย่าง (nonsampling bias) ได้แก่ กรอบและรายการ (listing and frame) ของตัวอย่าง การไม่ตอบคำถาม (nonresponse) ความคลาดเคลื่อนจากการวัด (measurement error) และความลำเอียงที่เกิดจากการเลือกตัวอย่าง (sampling bias) ได้แก่ ความลำเอียงจากการเลือกตัวอย่าง (selection bias) ความลำเอียงจากการประมาณการ (estimation bias) และความแปรปรวนจากการเลือกตัวอย่าง (sampling variability) ได้แก่ ขนาดตัวอย่าง (sample size) และความเป็นแบบเดียวกันของตัวอย่าง (sample homogeneity) (Henry, 1990: 34-46)

งานวิชาการที่เกี่ยวกับการวิเคราะห์ความคลาดเคลื่อนของผลการวิจัยด้านวิทยาศาสตร์ในบางสาขาวิชา เช่น ด้านการแพทย์ ด้านสุขภาพ เรียกว่า วิธีการวิเคราะห์ความลำเอียงในเชิงปริมาณ (quantitative bias analysis methods) และมีการศึกษากันมานานแล้วเช่นกัน โดยเฉพาะงานด้านการแพทย์มีปรากฏให้เห็นตั้งแต่ช่วงต้น ค.ศ. 1950 เกี่ยวกับการประเมินผลกระทบที่เกิดความผิดพลาดในการจำแนกประเภท (misclassification) และการประเมินผลกระทบที่เกิดจากปัจจัยรบกวนที่ไม่สามารถควบคุมได้ (uncontrolled confounding) โดยใช้แนวคิดหรือวิธีการแบบง่ายๆ ด้วยการใช้ตารางไขว้แบบ 2x2 วิเคราะห์เพื่อสรุปข้อมูลและข้อสันนิฐานที่เป็นไปได้เกี่ยวกับแหล่งที่มาของความลำเอียง (Lash & Others, 2016)

ความคลาดเคลื่อนที่มาจากความเข้าใจผิดหรือทำผิด (mistake/blunder) ในการวัด (measurement) หรือการคำนวณ (computation) แนวคิดเกี่ยวกับการศึกษาความคลาดเคลื่อนในการวัดมีความหลากหลายและไม่เหมือนกันในแต่ละสาขาวิชา บางงานเลือกศึกษาเฉพาะความคลาดเคลื่อนเดียว บางงานศึกษาหลายความคลาดเคลื่อน บางงานศึกษาวิธีการขจัดความคลาดเคลื่อน และบางงานอาจเน้นการวัดผลกระทบของผลการศึกษา แต่เนื้อแท้ของมุมมองที่แตกต่างกันต่างมีจุดหมายร่วมกัน คือ การศึกษาความคลาดเคลื่อนที่เกิดขึ้นภายในระเบียบวิธี (Biemer & others, 1991: 1; Bevington & Robinson, 1992: 1, 38-40)

โดยทั่วไปตำราที่เกี่ยวกับทฤษฎีความคลาดเคลื่อนเชื่อว่า ความคลาดเคลื่อนเป็นผลมาจากการกระจาย (dispersion) และความเบี่ยงเบน (deviation) ของข้อมูลที่ได้จากวัด และมีการจำแนกประเภทของความคลาดเคลื่อนเป็น 2 ประเภท ดังนี้ (Lash, Fox, & Fink, 2009: 13)

1. ความคลาดเคลื่อนเชิงสุ่ม (Random error) หรือความคลาดเคลื่อนเชิงสถิติ (statistical error) เป็นความคลาดเคลื่อนที่เกิดจากข้อมูลที่ได้มาจากการเลือกตัวอย่าง (sampling) โดยมีดัชนีความแม่นยำหรือความถูกต้อง (precision/accuracy)[4] เป็นค่าประมาณการของความคลาดเคลื่อน

2. ความคลาดเคลื่อนเชิงระบบ (Systematic error) เป็นความคลาดเคลื่อนที่เกิดจากเครื่องมือวัด (instrument) ที่ไม่มีคุณภาพหรือขาดความน่าเชื่อถือ โดยมีดัชนีความเที่ยงตรงเป็นค่าประมาณการของความคลาดเคลื่อน

การศึกษาเกี่ยวกับความคลาดเคลื่อนของการวิจัยจึงแบ่งได้เป็น 2 กลุ่มหลักตามทฤษฎีของความคลาดเคลื่อนดังนี้ (Biemer & others, 1991: 487)

1. กลุ่มที่เชื่อว่า ความคลาดเคลื่อนเกิดมาจากการเลือกตัวอย่าง เน้นการอธิบายความแปรปรวน-ความแปรปรวนร่วม (variance-covariance) ในเชิงโครงสร้างของประชากรทั้งหมด โดยมีสมมติฐานว่าความคลาดเคลื่อนเกิดมาจากข้อมูลหรือคำตอบที่ได้มาจากการเลือกตัวอย่างแบบสุ่มทั้งจากประชากรที่ทราบจำนวน (finite population) และประชากรที่ไม่ทราบจำนวน (infinite population) แม้ว่าความคลาดเคลื่อนจะเกิดมาจากทั้งการเลือกตัวอย่างแบบสุ่มทั้งจากประชากรที่ทราบจำนวน และประชากรที่ไม่ทราบจำนวน แต่ในวงวิชาการส่วนใหญ่มีความเชื่อและพูดถึงกันอย่างแพร่หลายว่า ตัวอย่างที่สุ่มมาจากประชากรที่ไม่ทราบจำนวนข้อมูลมีความคลาดเคลื่อน ซึ่งไม่เหมือนกับตัวอย่างที่สุ่มได้มาจากประชากรที่ทราบจำนวนที่มีความสงสัยและสนใจศึกษากันอย่างแพร่หลายเกี่ยวกับผลกระทบของความคลาดเคลื่อน (Burdick & Watrin, 2008)

[4] Precision ใช้กับผลของการวัดที่ใกล้กับค่าที่กำหนดไว้ (determined) ส่วน Accuracy ใช้กับผลของการวัดที่ใกล้กับค่าที่เป็นจริง (true value) (Bevington & Robinson 1992: 2)

2. กลุ่มที่เชื่อว่า ความคลาดเคลื่อนเกิดมาจากเครื่องมือวัด เน้นศึกษาความสัมพันธ์ของคำตอบที่มีความหลากหลายที่ได้มาจากตัวอย่างทั้งหมด และประเมินความถูกต้องหรือความคลาดเคลื่อนของคำตอบที่ได้มาจากแต่ละบุคคล

นอกจากที่กล่าวมา ตำราบางเล่มมีการแบ่งประเภทความคลาดเคลื่อนเพิ่มอีก 1 ประเภท คือ ความคลาดเคลื่อนเชิงบุคคล (gross error หรือ human error) ที่เกิดจากความเข้าใจผิดในการสร้างและใช้เครื่องมือวัดของนักวิจัย ได้แก่ ความไม่ตั้งใจจากการไม่ระมัดระวังในการตรวจสอบและตรวจซ้ำ ไม่รู้เท่าทันจากการละเลยหรือไม่เอาใจใส่ และจงใจเลือกข้อมูลแบบเจาะจง (Berendsen, 2011: 18-19)

ทฤษฎีการวัดแบบเก่ามีมุมมองว่า ความคลาดเคลื่อนเชิงสุ่มและความคลาดเคลื่อนเชิงระบบไม่สามารถนำมาสังเคราะห์ (synthesized) ร่วมกันได้ การประเมินความคลาดเคลื่อนที่เกิดขึ้นจากผลการวัดในขั้นสุดท้ายจากค่าความถูกต้อง (accuracy) ที่แสดงถึงความแม่นยำ (precision) และความน่าเชื่อถือ ถูกมองว่าเป็นมุมมองที่แคบ กล่าวคือ ความคลาดเคลื่อนเชิงสุ่มไม่ทำให้เกิดความลำเอียง (unbiased) แต่ความคลาดเคลื่อนเชิงระบบทำให้เกิดความลำเอียง (biased) ความคลาดเคลื่อนเชิงสุ่มเป็นต้นน้ำของการวัด (upstream measurement) ที่ทำให้เกิดการแจกแจงเชิงสุ่ม (random distribution) แต่ความคลาดเคลื่อนเชิงระบบเป็นปลายน้ำของการวัด (downstream measurement) ไม่ทำให้เกิดการแจกแจงเชิงสุ่ม จึงมีข้อเสนอว่า ทุกความคลาดเคลื่อนทำให้เกิดความลำเอียง ทุกความคลาดเคลื่อนทำให้เกิดการแจกแจงเชิงสุ่มได้ และมีผลกระทบทั้งความคลาดเคลื่อนเชิงสุ่มและความคลาดเคลื่อนเชิงระบบ ดังนั้นจึงต้องมองจากทั้งต้นน้ำและปลายน้ำ ประเมินความคลาดเคลื่อนด้วยค่าความไม่แน่นอน (uncertainty) ที่เบี่ยงเบนไปจากค่าที่เป็นจริงทั้งจากความคลาดเคลื่อนเชิงสุ่มและความคลาดเคลื่อนเชิงระบบ (Ye, Xiao, Shi, & Ling, 2016)

ปัจจุบันแนวคิดในการศึกษาความคลาดเคลื่อนในการวิจัยที่เริ่มจากการใช้ตารางไขว้แบบ 2x2 มีการพัฒนาไปใช้หลากหลายแนวคิดตามวัตถุประสงค์ของการศึกษา เช่น แบบจำลองมอนติคาร์โล (Monte

Carlo simulation methods)[5] วิธีการแบบเบส์ (Bayesian methods)[6] วิธีเชิงประจักษ์ (Empirical methods) และวิธีข้อมูลขาดหาย (Missing data methods) แต่ทุกวิธีมีแนวคิดพื้นฐานที่เหมือนกันดังนี้ (Lash & Others, 2016)

1. การหาสาเหตุ (Sources) ของการเกิดความคลาดเคลื่อนอย่างเป็นระบบ (systematic error) ได้แก่ การไม่สามารถควบคุมปัจจัยรบกวน (uncontrolled confounding) ความลำเอียงในการเลือก (selection biases) และความลำเอียงของสารสนเทศ (information biases)

2. ความลำเอียง (Biases) ของข้อมูลที่ได้มาจากการสังเกต (observed data) ที่ทำให้เกิดตัวแบบที่ลำเอียง (bias models)

3. การคำนวณ (Quantify) เกี่ยวกับทิศทาง (direction) ขนาด (magnitude) และความไม่แน่นอน (uncertainty) ที่เกิดจากความลำเอียงในการกำหนดค่า (value) ให้กับพารามิเตอร์

4. การตีความ (Interpreting) การสังเคราะห์ (syntheses) วิจารณ์ (critiques) และอ้างอิง (inferences) ผลลัพธ์อย่างผิวเผินหรือขัดแย้งกับความเป็นจริงตามธรรมชาติหรือในสังคมโดยเชื่อมั่นในผลการวิจัยเกินไป

ปัจจัยสำคัญที่ทำให้เกิดความเสี่ยงของการเกิดความคลาดเคลื่อน (risk of errors) คือ ข้อมูลระดับการวัด เวลาและค่าใช้จ่าย ขอบเขตของการวิจัย การวิเคราะห์ข้อมูล จำนวนประชากร ขนาดตัวอย่าง และวิธีการทางสถิติ ลักษณะของข้อมูล (body of data) ชุดหนึ่งอาจเหมาะสมกับเรื่องหนึ่งแต่อาจไม่เหมาะกับอีกเรื่องหนึ่ง ข้อมูลบางชุดที่ขาดหายอาจมีผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนโดยตรง ระดับการวัดอาจจำแนกประเภทกลุ่มตัวอย่างไม่ถูกต้อง เวลาหรือค่าใช้จ่ายอาจทำให้การออกแบบการวิจัยมีความละเอียดละออไม่เหมือนกัน ผู้ตอบคำถามในการวิจัยขนาดใหญ่ยากที่จะอยู่ภายใต้

[5] ใช้หลักการแจกแจงความน่าจะเป็น (probability distribution) แบบผกผันจากตัวอย่างแบบสุ่ม (random samples) ด้วยวิธีของเบส์ (Bayesian methods) จากเหตุการณ์ก่อนหน้าและเหตุการณ์ที่เกิดขึ้นหลัง

[6] ให้หลักการประมาณค่าพารามิเตอร์จากช่วงของความน่าเชื่อถือ (credibility interval) ของค่าพารามิเตอร์โดยไม่สนใจในการกำหนดสมมติฐานและทดสอบสมมติฐาน

สภาพแวดล้อมเหมือนกัน วิธีการทดสอบนัยสำคัญทางสถิติ (significance test) ที่ใช้ในการคาดคะเน (estimate) มีหลายมาตรฐาน มาตรฐานหนึ่งใช้การทดสอบนัยสำคัญทางสถิติของความถูกต้อง (valid) หรือที่เรียกว่า การทดสอบความเสี่ยงในการยืนยันการปฏิเสธความแตกต่างระหว่าง 2 กลุ่มประชาการ (type I error) ส่วนอีกมาตรฐานหนึ่งใช้การทดสอบนัยสำคัญทางสถิติของอำนาจจำแนกทางสถิติ (powerful) หรือที่เรียกว่า การตรวจสอบความแตกต่างระหว่าง 2 กลุ่มประชากร แต่ละวิธีการทดสอบมีการใช้จำนวนสัดส่วนของข้อมูลที่แตกต่างกัน (Bross, 1954)

ไทสัน โฮล์มส์ (Holmes, 2004) แบ่งความคลาดเคลื่อนเชิงสถิติที่เกิดจากความลำเอียง (bias) และความไม่แม่นจำ (imprecision) เป็นดังนี้

1. ความลำเอียงในการเลือกตัวอย่าง (Sampling bias) การเลือกตัวอย่างเริ่มต้นจากการกำหนดลักษณะเฉพาะของประชากรที่สนใจศึกษาและเป็นประชากรที่ถูกต้อง วิธีการลดความลำเอียงในการเลือกตัวอย่างให้น้อยที่สุด คือ การระบุประชากรให้ชัดเจน เพราะการกำหนดประชากรที่สนใจศึกษากว้างๆ เช่น มนุษย์ ธรรมชาติ สังคม รับรองได้ว่ามีความลำเอียงในการเลือกตัวอย่าง การเลือกสุ่มต้องรู้ความน่าจะเป็น (probability) ไม่เลือกตัวอย่างจากประชากรกลุ่มเดียวถ้าเลือกตัวอย่างจากประชากรมากกว่า 1 กลุ่ม ต้องใช้การเลือกตัวอย่างแบบช่วงชั้น (stratified sampling)

2. ความไม่แม่นยำในการเลือกตัวอย่าง (Sampling imprecision) ค่าเฉลี่ยของกลุ่มตัวอย่าง (sample mean) ในการวิจัยแต่ละครั้งจะมีระดับความแตกต่างกันตามไปตามขนาดของกลุ่มตัวอย่าง โดยทั่วไประดับความไม่เที่ยงตรงในการสุ่มสามารถดูได้จากค่าเฉลี่ยประชากรที่ประมาณการได้จากความเบี่ยงเบนมาตรฐาน (standard deviation) เช่น ความคลาดเคลื่อนมาตรฐาน (standard error) โดยความคลาดเคลื่อนมาตรฐานจะผันแปรไปตรงกันข้ามกับขนาดของกลุ่มตัวอย่าง หรืออีกนัยหนึ่งคือ ขนาดกลุ่มตัวอย่างจะมีความสัมพันธ์กับความแม่นยำ (precision) กล่าวคือ ความเที่ยงตรงจะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ความคลุมเครือนอกจากเกิดมาจากขนาดกลุ่มตัวอย่างแล้วยังอาจเกิดมาจากการเลือกตัวอย่างจากหลายแหล่งและไม่ได้ออกแบบการเลือกตัวอย่างแบบช่วงชั้น หากไม่ได้ทำการเลือกตัวอย่างแบบช่วงชั้น อาจต้องทำการเลือกตัวอย่างแบบช่วงชั้นหลังจากเก็บข้อมูล (poststratification) แต่อาจทำให้บางช่วงชั้น (strata) มีขนาดกลุ่มตัวอย่างจำนวนน้อย ซึ่งทำให้ความถูกต้องและอำนาจจำแนกเชิงสถิติลดลง ดังนั้นในการออกแบบการเลือกตัวอย่างนักวิจัยต้องกำหนดปัจจัยที่มีผลกระทบอย่างมีนัยสำคัญต่อ

ผลลัพธ์ให้ชัดเจน กำหนดขนาดกลุ่มตัวอย่างให้มีช่วงชั้นที่เท่ากันหรือใกล้เคียงกันในการทดสอบความแตกต่างระหว่างกลุ่มหรือช่วงชั้น และหากมีกลุ่มตัวอย่างขนาดเล็กก็ต้องทำการทดสอบซ้ำหลายๆครั้ง

3. ความลำเอียงในการวัด (Measurement bias) ความลำเอียงในการวัดสามารถเกิดขึ้นได้ตลอดเวลาในการวิจัย เช่น การเปลี่ยนอุปกรณ์ แหล่งข้อมูล ดังนั้นอาจต้องทำการศึกษาและเปรียบเทียบผลลัพธ์ภายใต้เงื่อนไขที่แตกต่างกันเพื่อตรวจสอบความสมมูล (equivalent) ของผลลัพธ์ โดยมีขอบเขตบนและล่าง (lower and upper bounds) ของค่าเฉลี่ยอยู่ระหว่าง ±1%

4. ความไม่แม่นยำในการวัด (Measurement imprecision) เกิดจากการใช้วิธีการวัดซ้ำ (technical repeats) โดยการวัดแบบเดียวกันซ้ำหลายๆครั้งจากสิ่งที่ทำการศึกษา ความคลาดเคลื่อนเชิงเทคนิค (technical error) ที่บางครั้งเรียกว่า การแปรผันภายในการวิจัย (intra-assay variation) ที่ปกติไม่มีการรายงานไว้ในผลการวิจัย (ค่าสัมประสิทธิ์ของการแปรผัน (coefficient of variation: C.V.) ที่ได้จากสัดส่วนของความเบี่ยงเบนมาตรฐาน (standard deviation) กับค่าเฉลี่ยของกลุ่มตัวอย่าง (sample means) ที่ควรจะมีค่าน้อยกว่า 5-10% โดยเฉพาะในกรณีที่กลุ่มตัวอย่างที่มีขนาดเล็กจะมีความลำเอียง (bias) มากขึ้นและควรทำการแก้ไข)

5. ความลำเอียงในการประมาณการ (Estimation bias) เป็นความคลาดเคลื่อนที่มีแนวโน้มที่สอดคล้องกัน (consistent tendency) และเป็นไปในทางเดียวกัน (direction) ของค่าพารามิเตอร์ที่เป็นลักษณะของประชากรที่สนใจจากการประมาณการด้วยกลุ่มตัวอย่าง เช่น ค่าเฉลี่ย ค่าความแปรปรวน ความลำเอียงในการประมาณการต่างไปจากความลำเอียงในการเลือกตัวอย่างและความลำเอียงในการวัด กล่าวคือ ไม่สนใจกับความลำเอียงที่เกิดจากการรวบรวมข้อมูลแต่เน้นความลำเอียงที่เกิดจากกระบวนการคำนวณ (computational process) มี 2 สาเหตุ ดังนี้

5.1 การไม่ให้ข้อมูลที่เกี่ยวกับข้อมูลขาดหาย (missing data) เช่น รายละเอียดเกี่ยวกับข้อมูลขาดหาย (informative missing data) ได้แก่ ข้อมูลขาดหายแบบสุ่ม (missing at random) ข้อมูลขาดหายที่ต้องปรับแก้ (nonignorable) ข้อมูลขาดหายแบบเป็นไปในทิศทางเดียวกัน (monotone missingness) และสิ่งที่ไม่ใช่รายละเอียดเกี่ยวกับข้อมูลขาดหาย (noninformative missing data) ได้แก่ ข้อมูลขาดหายแบบไม่ต้องปรับแก้ (ignorable)

5.2 การออกแบบไม่สมบูรณ์ (incomplete design) โดยไม่ทำการออกแบบไขว้กลุ่ม (crossover design) ทำให้ได้ผลการศึกษาที่ไม่ครอบคลุมปัจจัยเหตุ

6. ความไม่แม่นยำในการประมาณการ (Estimation imprecision) มีทั้งเหมือนและไม่เหมือน ความลำเอียงในการประมาณการ ที่เหมือนคือ ความคลุมเครือในการประมาณการเป็นความคลาดเคลื่อนในการประมาณการค่าของพารามิเตอร์ ที่ไม่เหมือนคือ มีความคลาดเคลื่อแบบสุ่ม (noise/random error) ความไม่เสถียร (instability) หรือความไม่สามารถเชื่อถือได้ (unreliability) ในการประมาณการ ซึ่งมีสาเหตุมาจากกลุ่มตัวอย่างมีขนาดเล็ก (small simple size) และกระบวนการเลือกตัวอย่าง (sampling process) ที่มีผลทำให้ลดประสิทธิภาพในการประมาณการ เพราะความถูกต้องที่เพิ่มขึ้นได้มาจากข้อมูลของขนาดของกลุ่มตัวอย่าง ซึ่งการลดขนาดของตัวอย่างเกิดขึ้นได้ทั้งในการวิเคราะห์และก่อนการประมาณการ

7. ความลำเอียงในการทดสอบสมมติฐาน (Bias in hypothesis testing) ผลการวิจัยไม่เคยมีความแน่นอน ต้องยอมรับว่าการทดสอบสมมติฐานด้วยวิธีการทางสถิติมีความไม่แน่นอน เพราะเป็นการค้นพบจากความน่าจะเป็นทางสถิติที่ได้จากการสิ่งที่สังเกต ความคลาดเคลื่อนแบบที่ 1 (type I error) หรือความคลาดเคลื่อนที่เกิดจากความน่าจะเป็นของความไม่ถูกต้องในการปฏิเสธสมมติฐานหลัก (null hypothesis: H_0) ที่สามารถเกิดได้จากหลายสาเหตุ ดังนี้

7.1 ระดับของการวัด (level of measurement) ความคลาดเคลื่อนแบบที่ 1 ประมาณร้อยละ 5 เกิดมาจากข้อมูลที่มีระดับการวัดแบบนามมาตรา (nominal level)

7.2 ข้อมูลที่คลุมเครือ (data snooping) มีผลทำให้เกิดความคลาดเคลื่อนแบบที่ 1 มากการข้อมูลที่มีระดับการวัดแบบนามมาตรา (nominal level)

7.3 การเลือกวิธีการทดสอบ (alternative testing) ระหว่างการทดสอบแบบหางเดียว (one-tailed) และการทดสอบแบบสองหาง (two-tailed) เพราะการทดสอบแบบหางเดียวกำหนดให้ใช้กับการทดสอบค่าเฉลี่ย (mean) ของประชากรกลุ่มหนึ่งที่มีขนาดใหญ่กว่าอีกกลุ่มหนึ่ง ส่วนการทดสอบแบบสองหางไม่กำหนดว่าความแตกต่างของค่าเฉลี่ยของประชากรกลุ่มหนึ่งต้องมากกว่าอีกกลุ่มหนึ่ง เนื่องจากการทดสอบแบบหางเดียวมีอำนาจจำแนกเชิงสถิติมากกว่า จึงทำให้เกิดความน่าจะเป็นของความถูกต้องในการปฏิเสธสมมติฐานหลัก

7.4 การทดสอบหลายสมมติฐาน (multiple hypothesis testing) การทดสอบแต่ละการทดสอบย่อมทำให้เกิดความคลาดเคลื่อนแบบที่ 1 เสมอ ดังนั้นเมื่อทำการทดสอบหลายสมมติฐานแบบไขว้จึงทำให้เกิดการผสมรวมของความคลาดเคลื่อนแบบที่ 1 โดยรวมตามมา

7.5 การกระจายของการเลือกตัวอย่าง (sampling distribution) ส่วนใหญ่เชื่อว่า การทดสอบตัวอย่างสองกลุ่มแบบที (*t*-test) ประชากรแต่ละกลุ่มไม่จำเป็นต้องมีการแจกแจงแบบปกติ (normally distribution) โดยทึกทักเอาว่า กลุ่มตัวอย่างจำนวนหนึ่งที่ดึงมาจากประชากรแต่ละกลุ่มและคำนวณค่าเฉลี่ยของตัวอย่างที่แตกต่างกัน ดังนั้นเมื่อดึงตัวอย่างมาจำนวนหนึ่งจากประชากรแต่ละกลุ่มและคำนวณสองค่าเฉลี่ยของกลุ่มตัวอย่างใหม่ด้วยวิธีการที่แตกต่างกัน กระบวนการทำซ้ำหลายๆครั้งจะทำให้เกิดการแจกแจงอย่างสมบูรณ์ของค่าเฉลี่ยของกลุ่มตัวอย่างที่แตกต่างกัน และมันเป็นการแจกแจงของการเลือกตัวอย่างที่เรามีข้อสันนิฐานว่า ปกติในการทดสอบตัวอย่างสองกลุ่มแบบที สามารถอธิบายได้ด้วยคุณลักษณะทางสถิติตามทฤษฎีแนวโน้มเข้าสู่ส่วนกลาง (The Central Limit Theorem: CLT) เพราะการแจกแจงของค่าเฉลี่ยของตัวอย่างมีแนวโน้มใกล้เคียงปกติเมื่อมีจำนวนขนาดของตัวอย่างเพิ่มมากขึ้น อย่างไรก็ตามเมื่อสองกลุ่มมีตัวอย่างและประชากรมีขนาดเล็ก มีความเบ้มาก (strongly skewed) มีหลายฐานนิยม (multiple mode) การกระจายของการเลือกตัวอย่างของความแตกต่างของค่าเฉลี่ยอาจจะไม่ปกติ (nonnormal) การใช้การทดสอบแบบทีอาจนำไปสู่การเกิดความคลาดเคลื่อนแบบที่ 1

นักวิจัยต้องรู้ว่าอะไรคือข้อสันนิฐาน (assumptions) ที่อยู่ภายใต้วิธีการทดสอบสมมติฐาน อย่างน้อยต้องทำการประเมินข้อมูล โดยทั่วไปกลุ่มตัวอย่างขนาดเล็กมีข้อจำกัดในการวิเคราะห์เพื่อตรวจสอบให้สอดคล้องกับข้อสมมุติ จำนวนประชากรที่เป็นตัวอย่างเท่ากับหรือมากกว่า 30 ตัวอย่าง เป็นกฎคร่าวๆที่อยู่ภายใต้ข้อสมมุติในหลายการทดสอบเชิงสถิติ แม้กระทั่งการทดสอบแบบไม่มีพารามิเตอร์ (nonparametric test) ก็มีการกำหนดข้อสมมุติเกี่ยวกับประชาการที่ได้มาจากการสุ่ม ดังนั้นในการวิจัยควรจะแสดงรายงานสถิติผลการตรวจสอบข้อสมมุติ การเลือกวิธีการที่เหมาะสม (appropriate methods) สำหรับข้อมูลที่เหมาะสม (fitting data) และทำการทดสอบสมมติฐานอย่างสมดุล วิธีการต่างๆขึ้นอยู่กับข้อสมมุติที่เข้มแข็งมากสามารถทำให้เกิดความแข็งแกร่งเมื่อข้อมูลตรงกับข้อสมมุติอย่างชัดเจน แต่หากฝ่าฝืนข้อสมมุติจะทำให้เกิดความคลาดเคลื่อนแบบที่ 1

8. ความไม่แม่นยำในการทดสอบสมมติฐาน (Imprecision in hypothesis testing) เป็นความคลุมเครือในการวัดที่เกิดความคลาดเคลื่อนแบบที่ 2 ในการทดสอบสมมุติฐาน คือ ปฏิเสธ (reject) สมมติฐานทางเลือก (alternative hypothesis) ที่เป็นจริง (true) ซึ่งมีความเป็นไปได้ว่าเกิดจากอำนาจจำแนกของสถิติ (statistical power) ถ้าความคลาดเคลื่อนแบบที่ 2 เกิดจาก β อำนาจจำแนกของสถิติที่ได้จาก 1-β อำนาจจำแนกดังกล่าวเป็นไปได้ที่จะปฏิเสธสมมติฐานหลักเมื่อสมมติฐานทางเลือกเป็นจริง ความคลาดเคลื่อนแบบที่ 2 สามารถเกิดจากหลายแหล่งข้อมูล (sources) ที่มีความคลุมเครือ ซึ่งเกิดขึ้นในกระบวนการทดสอบสมมติฐาน ดังนั้น ความคลาดเคลื่อนแบบที่ 2 จึงเกิดเพิ่มมากขึ้นโดย 1) กลุ่มตัวอย่างมีขนาดเล็กกว่า (smaller sample sizes) 2) การใช้เทคนิคซ้ำ (larger technical error) หรือ 3) ตัวประมาณการมีประสิทธิภาพต่ำ (less efficient estimators) นอกจากนั้น การใช้ขนาดตัวอย่างที่มีจำนวนไม่เท่ากันในการเปรียบเทียบระหว่างกลุ่มก็เป็นสาเหตุทำให้เสียโอกาสในการเพิ่มอำนาจจำแนกของสถิติ

9. ความลำเอียงในการรายงาน (Reporting bias) เป็นความลำเอียงที่อาจไม่สามารถตรวจสอบได้ (undetectable) ที่เกิดจากการไม่ตีพิมพ์ผลงานที่แสดงผลลบหรือไม่ชัดเจน (file drawer problem) เช่น ผู้เขียนหรือบรรณาธิการเลือกที่จะไม่ตีพิมพ์ข้อค้นพบที่ไม่มีนัยสำคัญทางสถิติ (insignificance) ที่เป็นความจริง เพราะการศึกษามีขนาดตัวอย่างไม่เพียงพอที่จะนำมาอ้างอิงทางสถิติ (underpowered studies) อย่างไรก็ตามแม้ว่าผลการศึกษาจะไม่มีนัยสำคัญทางสถิติแต่ก็ให้ร่องรอยเชิงประจักษ์ (empirical suggestion) และยอมรับว่ามีจุดอ่อนในการสรุปสมมติฐาน แต่ก็มีความสำคัญและเป็นประโยชน์ในการทบทวนวรรณกรรม และสามารถนำไปทำการทดสอบแบบเข้มงวดในเวลาต่อมาด้วยการใช้ขนาดตัวอย่างขนาดใหญ่ (fully powered studies)

10. ความไม่แม่นยำในการรายงาน (Reporting imprecision) เหมือนกับการไม่ตีพิมพ์ผลงานที่แสดงผลลบหรือไม่ชัดเจน แต่เกิดจากปิดบังผู้อ่านที่ไม่เข้าใจ (misunderstanding) ผลการวิจัย เช่น การวิจัยรายงานค่าเฉลี่ยของกลุ่มตัวอย่างเป็นค่า ± แต่ไม่แสดงค่าเบี่ยงเบนมาตรฐาน ความคลาดเคลื่อนมาตรฐาน หรือค่าการกระจายอื่นๆ ของอีกค่าหนึ่ง เพราะค่าพารามิเตอร์การกระจาย (dispersion parameters) มีความสำคัญกับนักวิจัยคนอื่นๆที่จะนำไปใช้ในการวางแผนกำหนดขนาดตัวอย่างของงานวิจัยในอนาคต หรือมีการแสดงระดับนัยสำคัญของสถิติ (*P* value) แต่ไม่มีข้อมูลเกี่ยวกับวิธีการทางสถิติที่ใช้

ความคลาดเคลื่อนจากเครื่องมือวัดและการวัด

การวัดเป็นที่บรรจบกันของทฤษฎีที่เป็นนามธรรม (abstract theory) และปรากฏการณ์เชิงประจักษ์ (empirical phenomena) ที่ต้องการค้นคว้าหาคำตอบเพื่อนำไปสู่การอธิบาย และการวัดเป็นการกำหนดตัวเลขให้กับคำตอบที่ได้จากบุคคล การทดสอบ หรือแบบสอบถาม (Traub, 1994; Thye, 2000)

ความเสี่ยงของการเกิดความคลาดเคลื่อนของผลการวิจัยเป็นโอกาสในการเกิดความคลาดเคลื่อนจากการวัด (measurement error) หรืออาจเรียกว่า ความคลาดเคลื่อนจากคำตอบ (response error) ที่เป็นผลมาจากความแตกต่างระหว่างค่าที่ได้จากการสังเกต (observed) หรือการคำนวณ (calculated) กับค่าที่เป็นจริง (true value) หรือผลของการวิเคราะห์ข้อมูลและการสรุปผลการวิจัยที่เบี่ยงเบน (deviations) ไปจากการสะท้อนถึงความเป็นจริง (true reflections) ของประชากร ซึ่งมีความเกี่ยวข้องกับหลายสาเหตุ แต่ที่พบมากที่สุด คือ ความคลาดเคลื่อนที่เกิดจากเครื่องมือ (instrument error) และความคลาดเคลื่อนที่เกิดการเลือกตัวอย่าง (sampling error) (Biemer & others, 1991: 1, 618; Bevington & Robinson, 1992: 1; Buonaccorsi, 2010: 1)

จากบทความเรื่อง “ความคลาดเคลื่อนในการสำรวจ (On Errors in Surveys)” ของวิลเลียม เอ็ดเวิดส์ เดมมิ่ง (William Edwards Deming) น่าจะเป็นหลักฐานเชิงประจักษ์ที่บอกได้ว่า การศึกษาเกี่ยวกับความคลาดเคลื่อนในการวิจัยด้านสังคมศาสตร์มีจุดเริ่มต้นในปี ค.ศ. 1914 จากข้อมูลการสัมภาษณ์เพื่อศึกษาด้านสังคมของสำนักงานสถิติแห่งชาติของสหรัฐอเมริกา (United State Census Bureau) ที่ปรากฏอยู่ในงานของสจ๊วต เอ ไรซ์ (Struart A. Rice) เรื่อง “ความลำเอียงแพร่เชื้อได้ในการสัมภาษณ์ (Contagious Bias in the Interview)” ที่พบว่า ผู้ให้ข้อมูลให้เหตุผลของความยากจนแตกต่างกันไปตามการปรุงแต่งของผู้สัมภาษณ์ (Rice, 1929; Deming, 1944)

ทาย (Thye, 2000) สรุปว่า โดยทั่วไปความคลาดเคลื่อนในการวัดมี 2 ประเภท คือ ความคลาดเคลื่อนในการวัดแบบคงที่ (constant measurement error) เป็นความคลาดเคลื่อนที่เหมือนกันจากการสังเกตในการวัดแต่ครั้ง และมีผลต่อความเที่ยงตรงเชิงโครงสร้าง (construct validity) และความคลาด

เคลื่อนในการวัดแบบสุ่ม (random measurement error) เป็นความคลาดเคลื่อนที่ไม่เหมือนกันจากการสังเกตในการวัดแต่ละครั้ง และมีผลต่อความน่าเชื่อถือได้ (reliability) ซึ่งความคลาดเคลื่อทั้ง 2 ประเภทนั้นมีความสำคัญ แต่ความน่าเชื่อถือได้เป็นสิ่งที่ต้องให้ความสนใจเป็นอย่างยิ่งโดยเฉพาะในการวิจัยที่มีเป้าหมายในการพัฒนาทฤษฎี เพราะการลดความคลาดเคลื่อนแบบสุ่มจะช่วยให้อำนาจจำแนกเชิงสถิติดีขึ้นและความน่าเชื่อถือสูงขึ้น เมื่อการวัดมีความน่าเชื่อถือสูงจะทำให้มีเที่ยงตรงภายใน (internally valid) และให้ประโยชน์เชิงทฤษฎี (theoretically informative)

เป็นที่ทราบกันอย่างแพร่หลายว่า ความคลาดเคลื่อนที่เกิดจากการวัดหรือความคลาดเคลื่อนจากการสังเกตมีผลกระทบอย่างมากต่อการประมาณการและการอ้างอิงเชิงสถิติ อีกทั้งยังมีผลการศึกษาจำนวนมากที่ตรวจสอบพบว่า รายงานผลการศึกษาที่ใช้ข้อมูลจากการสำรวจมีความคลาดเคลื่อน (Bollinger, 1998; Schennach, 2012) และการวัดที่ขาดความน่าเชื่อถือเป็นสาเหตุทำให้ได้ค่าประมาณการที่ต่ำกว่าความเป็นจริงที่นำไปสู่การเพิ่มความเสี่ยงของการเกิดความคลาดเคลื่อนแบบที่ 2 (type II error) (Osborne, 2008: 239)

ความคลาดเคลื่อนในการค้นหาคำตอบ (error in inquiry) เกิดมาจากหลายสาเหตุ ดังนี้ (Babbie, 2013: 6-7)

1. ความไม่เที่ยงตรงในการสังเกต (Inaccurate observation) เช่น ไม่ตั้งใจสังเกต เลือกใช้เครื่องมือที่ไม่เหมาะสม

2. การสรุปเกินความจริง (Overgeneralization) อันเนื่องมาจากการสังเกตที่ถูกจำกัดขอบเขต (limited observation) ความกดดัน (pressure) เร่งรีบสรุปจากเรื่องบางส่วน สรุปในระดับหนึ่งเพื่อเอาตัวรอด (survive)

3. การเลือกสังเกต (Selective observation) การสรุปเกินความจริงนั้นสามารถนำไปสู่การเลือกสังเกต เรามีแนวโน้มที่จะเน้นสถานการณ์และเหตุการณ์ที่ตรงหรือเหมาะ (fit) กับแบบอย่าง (pattern) มีความลำเอียงในการเลือกเชิงเชื้อชาติ (racial) และชาติพันธุ์ (ethnic) บางครั้งเราอาจออกแบบการวิจัยเฉพาะจำนวนและประเภทที่เป็นประโยชน์จากการสังเกตเพื่อให้ได้ข้อสรุป เช่น เลือกกลุ่มคนที่มีแนวโน้มจะสนับสนุนให้มีการทำแท้งอย่างเสรี

4. การให้เหตุผลอย่างไม่มีตรรกะ (Illogical reasoning) ด้วยการกล่าวว่า กฎใดๆจะมีข้อยกเว้นเสมอ (the exception that proves the rule)

แฟรงค์ ชมิดท์และจอห์น ฮันเตอร์ (Schmidt & Hunter, 1996 cited by Thye, 2000) อธิบายว่า ความคลาดเคลื่อนในการวัด (measurement errors) มี 3 ประเภท ดังนี้

1. ความคลาดเคลื่อนแบบชั่วคราว (Transient errors) ที่เป็นผลมาจากคำตอบของปัจเจกบุคคล (individual responses) ที่ได้มาจากการวิจัยด้วยวิธีการเดียวกัน แต่มีการวิจัยในเวลาที่แตกต่างกัน ทุกครั้งที่มีการวิจัยใหม่และประชากรหรือกลุ่มตัวอย่างกลุ่มใหม่ ความไม่แน่นอนก็จะเกิดขึ้น ปัจจัยต่างๆสามารถมีอิทธิพลต่อพฤติกรรมของประชากร ความคลาดเคลื่อนแบบชั่วคราวมีสาเหตุมาจากการเปลี่ยนแปลงที่ผิดปกติของสิ่งแวดล้อม เช่น อุณหภูมิ ระดับความชื้น สภาพของแสง เสียง และเวลา และพฤติกรรมของผู้ถูกวิจัยและประชากร เช่น ระดับความสนใจ การยอมรับ แรงจูงใจ อารมณ์ และความเข้าใจ ปัจจัยเหล่านี้ทำให้เกิดความผันแปร (variability) ในข้อมูล ไม่ได้มีสาเหตุมาจากตัวแปรอิสระ (independent variable) ผลกระทบของความคลาดเคลื่อนแบบชั่วคราวจึงไม่จำเป็นต้องทำการวิจัยซ้ำใหม่และไม่จำเป็นต้องพิจารณาความคลาดเคลื่อนในการวัด

2. ความคลาดเคลื่อนแบบสุ่ม (Random errors) มีแนวคิดคล้ายกับความคลาดเคลื่อนแบบชั่วคราว แต่มีความแตกต่างกันมากในเรื่องของช่วงเวลาที่สั้นกว่า กล่าวคือ ความคลาดเคลื่อนแบบสุ่มเกิดความแปรปรวนภายใน (within) ในการวิจัยแต่ละครั้ง ส่วนความคลาดเคลื่อนแบบชั่วคราวเกิดจากความแปรปรวนระหว่าง (between) ในการวิจัยแต่ละครั้ง เป็นความคลาดเคลื่อนที่ไม่ได้มีสาเหตุมาจากตัวแปรอิสระ แต่มีผลกระทบต่อตัวแปรตามหรือผลลัพธ์ที่ได้จากการวัด จึงจำเป็นต้องพิจารณาความคลาดเคลื่อนในการวัด

3. ความคลาดเคลื่อนแบบเฉพาะกรณี (Specific errors) เป็นความคลาดเคลื่อนที่มีความสัมพันธ์กันเล็กน้อย มีสาเหตุมาจากทุกลักษณะที่ไม่เหมือนกัน (every unique) ได้แก่ ข้อคำถามในแบบสอบถาม (questionnaire item) พฤติกรรมที่วัด (behavioral measure) และบริบทของการวิจัย (experimental context) เพราะความจริงแต่ละสิ่งและทุกสิ่งที่ทำการวัดจะมีลักษณะเฉพาะของตัวมันเอง บางครั้งการวัดก็ไม่สามารถทำได้ ดังนั้นผลที่เกิดจึงมีความแปรปรวนในตัวของมันเอง แต่ละการวิจัยเพื่อยืนยันตามทฤษฎีจะ

ได้ผลลัพธ์ที่แตกต่างกัน ความสัมพันธ์ของความแปรปรวนจากปัจจัยเฉพาะในแต่ละการวัดไม่มีผลมากนักต่อการสร้างทฤษฎีทั่วไป แต่อย่างไรก็ตามควรที่จะพิจารณาความคลาดเคลื่อนในการวัด

นอกจากการวัดความน่าเชื่อถือด้วยวิธีการเชิงประจักษ์ ยังมีการพัฒนาทฤษฎีเชิงสาเหตุ (causal theories) เพื่อนำไปใช้กับปรากฏการณ์ของโลกที่เป็นจริง โดยใช้การค้นหาและนำไปสู่การสร้างทฤษฎีด้วยการตรวจสอบความเที่ยงตรงภายใน (internally valid) และความเที่ยงตรงภายนอก (external valid) ดังนี้ (Thye, 2000)

1. ความเที่ยงตรงภายใน (Internal validity) เป็นตรรกะของการทดสอบการวิจัยในการสร้างกฎเชิงสาเหตุ (causal laws) โดยการออกแบบการวิจัยอย่างรอบครอบในการกำหนดวิธีการวิจัยตัวแปรที่เป็นสาเหตุที่มีผลต่อการวัดตัวแปรอิสระ นักวิจัยต้องพยายามควบคุมสภาพแวดล้อม ความเป็นแบบเดียวกันของตัวอย่าง (homogeneous samples) ระเบียบวิธีการที่เป็นมาตรฐาน (standardized protocols) และเทคนิคการวิจัยแบบปกปิดทั้งนักวิจัยและผู้ถูกวิจัย (double blinds) ความคลาดเคลื่อนในการวัดไม่ได้แยกออกจากความสัมพันธ์เชิงสาเหตุอย่างแท้จริง แต่จะทำให้เกิดการเข้าใจผิดหรือเกิดการบิดเบือน (distort) ในข้อมูลของการวิจัยนั้น เพราะเป็นไปไม่ได้ที่จะสรุป (draws) ความถูกต้องในการอ้างอิงจากข้อมูลที่ลำเอียง การวัดความเที่ยงตรงจึงเป็นเงื่อนไขที่จำเป็นสำหรับความถูกต้องภายใน

2. ความเที่ยงตรงภายนอก (External validity) นักวิจัยกลุ่มหนึ่งเชื่อว่า การวิจัยต้องทำในสภาพแวดล้อมที่อยู่ในห้องทดลอง การสร้างความรู้ไม่สามารถทำได้ในสถานการณ์ที่เป็นธรรมชาติ ส่วนนักวิจัยอีกกลุ่มหนึ่งให้เหตุผลว่า การวิจัยในห้องปฏิบัติการมีเป้าหมายในการตรวจสอบความคิดเชิงทฤษฎีแบบบริสุทธิ์ (pure theoretical ideas) และเป็นไปไม่ได้ที่นักวิจัยจะจำลองเหตุการณ์ทางสังคม (social situations) ที่เกิดขึ้นตามธรรมชาติไปค้นคว้าในห้องทดลอง เป้าหมายของการวิจัย คือ การสร้างการวิจัยเพื่อหาสาเหตุใหม่ที่อยู่ภายในขอบเขตของทฤษฎี

ในปี ค.ศ. 1986 เอ็ดวิน เอ. ล็อค (Edwin A. Lock) ได้ทำการตีพิมพ์ผลการค้นคว้าของเขาและนักวิจัยคนอื่นๆ พบว่า มีความสอดคล้องกันอย่างชัดเจนระหว่างภาคสนาม (fields) กับข้อค้นพบจากการวิจัยในห้องปฏิบัติการ (labs) นักวิทยาศาสตร์ทุกสาขาวิชาจะทำการควบคุมภายในห้องทดลองเพื่อทดสอบการตรวจสอบสมมติฐานเชิงทฤษฎี (theoretical hypotheses) ส่วนนักสังคมศาสตร์จะใช้การวิจัยหรือ

การสำรวจที่คล้ายคลึงกันในเรื่องที่เกี่ยวกับเหตุการณ์ทางสังคมเหมือนกับที่นักวิทยาศาสตร์ธรรมชาติใช้เครื่องชั่ง คาน และเลนส์ ซึ่งคล้ายคลึงกับรายละเอียดที่มีอยู่ตามธรรมชาติ (natural items) ถ้าการค้นพบจากการวิจัยไม่สรุปไปสู่โลกที่เป็นประสบการณ์ในชีวิตประจำวันจะทำการวิจัยไปทำไม ตอบสั้นๆ คือ เป็นการวิจัยเพื่อทดสอบทฤษฎี การทดสอบทฤษฎีต่างๆมักนำไปใช้ประโยชน์โดยตรงนอกสภาพแวดล้อมของห้องปฏิบัติการ (Thye, 2000)

ความคลาดเคลื่อนจากประชากรและตัวอย่าง

ความคลาดเคลื่อนในการประมาณการ (estimate) เกิดขึ้นได้ทั้งจากความคลาดเคลื่อนในการเลือกตัวอย่าง (sampling errors) และความคลาดเคลื่อนที่ไม่ใช่การเลือกตัวอย่าง (nonsampling errors) โดยทั่วไปอาจเชื่อกันว่า การวัดตัวแปรที่ได้มาจากหน่วย (units) ของตัวอย่างไม่มีความคลาดเคลื่อน เพราะความคลาดเคลื่อนเป็นการประมาณการที่ได้มาจากประชากรส่วนหนึ่งที่รวมอยู่ในกลุ่มตัวอย่าง แต่ในความเป็นจริงข้อมูลที่ได้จากบางกลุ่มตัวอย่างอาจไม่ได้เป็นประชากรเป้าหมายหรืออยู่นอกกรอบของการเลือกตัวอย่าง (sampling frames) และการปฏิเสธในการตอบคำถามของกลุ่มตัวอย่างก็ทำให้ข้อมูลที่ได้มาจากตัวอย่างไม่เป็นตัวแทน (unrepresentative) ของประชากรและทำให้เกิดความลำเอียง (biased) ในการประมาณการได้เช่นกัน (Cochran, 1977: 359; Thompson, 2012: 5; Arnab, 2017: 3-4)

ความคลาดเคลื่อนในการสำรวจข้อมูลจากตัวอย่าง (sample surveys) หรือการเลือกตัวอย่าง (samplings) เกิดมาจากหลายสาเหตุ แต่แบ่งได้เป็น 2 กลุ่มใหญ่ คือ ความคลาดเคลื่อนที่ไม่ได้มีสาเหตุมาจากการสังเกต (error of nonobservations) ได้แก่ การเลือกตัวอย่าง (sampling) ความครอบคลุมหรือคุ้มรวม (coverage) และการไม่ตอบคำถาม และความคลาดเคลื่อนที่มีสาเหตุมาจากการสังเกต (error of observations) ได้แก่ ผู้ทำการสัมภาษณ์หรือผู้รวบรวมข้อมูล (interviewer/data collector) การตอบคำถาม (respondent) เครื่องมือวัด (instrument) และวิธีการรวบรวมข้อมูล (method of data collection)

การกำหนดขนาดตัวอย่าง (sample size) ในการคำนวณอำนาจจำแนกทางสถิติได้มาจาก 4 ปัจจัย คือ ความน่าจะเป็น (probability) ในการปฏิเสธสมมติฐานหลักถ้าสมมติฐานทางเลือกเป็นจริง

ขนาดอิทธิพล (effect size) สำหรับสมมติฐานทางเลือก ระดับอัลฟาหรือนัยสำคัญ (alpha or significance level) ซึ่งเป็นความน่าจะเป็นของการเกิดความคลาดเคลื่อนแบบที่ 1 ที่เกิดจากความไม่ถูกต้องในการปฏิเสธสมมติฐานหลักเมื่อสมมติฐานหลักเป็นจริง และรูปแบบการทดสอบทางสถิติเป็นแบบหนึ่งหรือสองด้าน (one or two sided)

โฮล์มส์ (Holmes, 2004) จำแนกแหล่งที่มาของความคลาดเคลื่อนเชิงสถิติที่เกิดจากความคลาดเคลื่อนในการเลือกตัวอย่าง (sampling errors) ไว้ 2 แบบ คือ

1. ความลำเอียง (Bias) ที่เป็นความคลาดเคลื่อนที่มีแนวโน้มที่สอดคล้องกัน (consistent tendency) และเป็นไปในทางเดียวกัน (direction) ซึ่งมีผลทำให้เกิดความลำเอียงในการประมาณการด้วยพารามิเตอร์ (parameters) ทุกประเภท ได้แก่ ตำแหน่ง (location) เช่น ค่าเฉลี่ย (mean) การกระจาย (dispersion) เช่น ความแปรปรวน ซึ่งสามารถทำให้เกิดความเข้าใจผิดจากความไม่เหมือนกันของประชากรที่เป็นจริง การประมาณการความแปรปรวนมากเกินไปจะมีผลทำให้สูญเสียอำนาจจำแนกเชิงสถิติ แต่หากการประมาณการความแปรปรวนน้อยเกินไปจะเพิ่มโอกาสให้เกิดความผิดพลาดในการสรุปผลที่ได้จากการวิจัย

2. ความไม่แม่นยำ (Imprecision) ของความคลาดเคลื่อนที่เกิดจากผลกระทบแบบไร้ทิศทาง (nondirectional noise) คือ กระบวนการเลือกตัวอย่างจากประชากรที่สนใจทำให้เกิดความคลาดเคลื่อนจากการเลือกตัวอย่าง เกิดความคลาดเคลื่อนในกระบวนการวัด และมีผลต่อเนื่องไปถึงกระบวนการประมวลผลข้อมูลที่ได้มาจากการรวบรวมข้อมูล เกิดความคลาดเคลื่อนในการประมาณการ (estimation) จากค่าพารามิเตอร์ไปสู่กลุ่มประชากรที่ใช้ในการวิจัย ผลของการวิจัยที่ไม่มีความแน่นอนจะนำไปสู่ความคลาดเคลื่อนแบบที่ 1 และความคลาดเคลื่อนแบบที่ 2 ในการทดสอบสมมติฐานและเมื่อการวิจัยเสร็จสมบูรณ์แล้วนักวิจัยเลือกที่จะไม่รายงานผลการทดสอบที่ไม่มีนัยสำคัญทางสถิติ (statistically insignificant)

ความคลาดเคลื่อนจากสถิติและการทดสอบสมมติฐาน

สถิติเป็นทั้งศาสตร์และศิลป์สำหรับใช้ในการอ้างอิงจากข้อมูลที่มีความไม่คงที่ (uncertain data) ไม่ใช่ใช้สำหรับค้นหา $p<0.05$ อย่างไร้สติ แต่คนส่วนใหญ่คิดว่าสถิติเป็นเรื่องของพีชคณิต (algebra) และ

การคำนวณ (computation) ซึ่งเป็นมุมมองที่แคบ และใช้การทดสอบนัยสำคัญทางสถิติในทางที่ผิด เช่น ใช้มากเกินไป (overuse) ใช้น้อยเกินไป (underuse) ใช้ผิด (misuse) แปลความหมายผิด (misinterpretation) และใช้เพื่อให้น่าเชื่อถือ (clear fishing around) (Rigby, 1998) ซึ่งสอดคล้องกับที่ชมิดท์และฮันแตอร์ (Schmidt & Hunter, 1989, 1996 cited by Thye, 2000) กล่าวไว้ว่า บางกรณีมีการใช้สูตรผิด หรือบางกรณีใช้สูตรถูกแต่ตีความหมายผิด

การประเมินระดับของการเกิดความคลาดเคลื่อนในการอ้างอิงเป็นสิ่งที่สำคัญเมื่อทำการวางแผนการทำวิจัยเชิงยืนยัน (confirmatory research) แม้ว่าความคลาดเคลื่อนในการอ้างอิงและการวิเคราะห์อำนาจจำแนกของสถิติจะเป็นหัวข้อมาตรฐานในหนังสือด้านสถิติมาอย่างช้านาน แต่โดยทั่วไปในการวิจัยเชิงสำรวจ (exploratory research) จะเน้นการแสดงค่าความน่าจะเป็น (*p* value) หรือขนาดอิทธิพล (effect size) และมีน้อยมากที่ให้ความสำคัญกับขนาดตัวอย่าง (sample size) อำนาจจำแนก (power) และความคลาดเคลื่อนในการอ้างอิง (inferential errors) และที่สำคัญในการออกแบบที่ดีในการวิจัยเชิงยืนยัน (confirmatory research) ต้องแสดงให้เห็นปฏิกิริยากระทบต่อกัน (interactions) ของขนาดอิทธิพล ขนาดกลุ่มตัวอย่าง นัยสำคัญทางสถิติ และความคลาดเคลื่อนในการอ้างอิง รวมถึงกระบวนการวิเคราะห์ทั้งหมดในการตัดสินใจตั้งแต่ก่อนทำการรวบรวมข้อมูลจนถึงการทำการรวบรวมข้อมูล ได้แก่ วิธีการทางสถิติ (statistical methods) เงื่อนไขสำหรับการยอมรับข้อพิสูจน์ (acceptable evidence) รายละเอียดในการปรับแก้ข้อมูล (data adjustment) และเงื่อนไขในการตัดทิ้งข้อมูล (excluding data) จากการวิเคราะห์ หากไม่สามารถแสดงรายละเอียดไว้ให้เห็นก่อนได้ การศึกษานั้นก็เป็นเพียงการสำรวจมากกว่าการยืนยัน ดังนั้นการตัดสินใจอย่างมีระเบียบวิธี (methodological decisions) จึงควรเปิดเผยหรือแสดงให้เห็นกระบวนการและเงื่อนไขทั้งหมดก่อนเริ่มการเก็บรวบรวมข้อมูล (Kennedy, 2015)

พูนาม ซิงห์, ราเจส ซิงห์ และคาร์ลอส โบว์ซา (Singh, Singh, & Bouza, 2016) ศึกษาผลกระทบของความคลาดเคลื่อนในการวัดและความคลาดเคลื่อนของการไม่ตอบคำถามหรือค่าสูญหาย (missing data) ในการประมาณการค่าเฉลี่ยของประชากร (population mean) โดยใช้ชุดข้อมูลที่แตกต่างกัน ระหว่างไม่มีความคลาดเคลื่อนในการวัดและมีความคลาดเคลื่อนในการวัดเพื่อเปรียบเทียบประสิทธิภาพ

ของตัวประมาณการ (efficiency of estimators) พบว่า ความคลาดเคลื่อนในการวัดและความคลาดเคลื่อนของการไม่ตอบคำถามมีผลกระทบต่อค่าร้อยละที่มีความสัมพันธ์กับประสิทธิภาพ (percentage relative efficiency) ของตัวประมาณการในอัตราที่สูง

การเลือกใช้สถิติแต่ละประเภทในการวิเคราะห์ข้อมูลมีข้อตกลง (assumptions) ที่กำหนดให้ต้องปฏิบัติตาม โดยทั่วไปในการเลือกใช้สถิติเพื่อทดสอบนัยสำคัญ นักวิจัยส่วนใหญ่มีความเห็นและเชื่อกันมาอย่างยาวนานว่า ไม่มีเหตุผลอื่นที่จะทำให้เกิดการละเมิด (violate) ข้อตกลงเกี่ยวกับการทดสอบสถิติแบบมีพารามิเตอร์ หากข้อมูลที่ใช้มีระดับการวัดเป็นช่วงมาตราหรือสัดส่วนมาตรา (Sheskin, 2007: 108-109)

แม้ว่าสถิติแบบมีพารามิเตอร์ส่วนใหญ่วางอยู่บนฐานของข้อมูลที่มีระดับการวัดตั้งแต่ช่วงมาตราขึ้นไปและข้อมูลต้องมีการแจกแจงแบบปกติแล้ว แต่สถิติแต่ละประเภทยังมีข้อตกลงเฉพาะเพิ่มเติมอีก เช่น การวิเคราะห์เปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มตัวอย่าง ข้อมูลแต่ละกลุ่มตัวอย่างต้องมาจากประชากรที่มีความเป็นแบบเดียวกันของความแปรปรวน (homogeneity of variance) การวิเคราะห์ค่าเฉลี่ยแบบทดสอบซ้ำ ข้อมูลของแต่ละตัวอย่างต้องมีความเป็นอิสระจากกัน (independence) และการวิเคราะห์ตัวแบบการทำนาย ข้อมูลระหว่างตัวแปรที่ใช้ทำนาย (predictor variable) และตัวแปรที่ถูกทำนาย (predicted variable) ต้องมีความสัมพันธ์กันเชิงเส้นตรง (linearly) หากไม่ให้ความสนใจข้อตกลงเหล่านี้ ผลการวิเคราะห์จะไม่มีความถูกต้องและทำให้ได้ข้อสรุปที่ผิด (Field, 2009: 132-133, 2013: 165-176)

ความคลาดเคลื่อนจากการสรุปผลการวิจัย

บาร์บารา เอฟ. มีเกอร์ และโรเบิร์ต ลิค (Barbara F. Meeker และ Robert Leik 1995: 640 cited by Thye, 2000) กล่าวว่า การสร้างข้อสรุปเชิงทฤษฎี (theoretical generalization) เป็นการความเชื่อมโยงระหว่าง ความรู้เชิงทฤษฎี (theoretical knowledge) กับการนำไปใช้กับปัญหาต่างๆ (applied problems) ข้อค้นพบจากการวิจัยต้องทำการตีความหรือแปลผลด้วยตาไปสู่ผลการค้นพบที่ทำการศึกษามาแล้วในอดีตและการประยุกต์ใช้กับสถานการ์ในปัจจุบัน

การสำรวจ (exploration) และการยืนยัน (confirmation) เป็นสิ่งจำเป็นทั้งคู่สำหรับการหาความรู้ที่ได้จากการสังเกต การค้นคว้า และการวิจัย ปัญหาหนึ่งที่นักวิจัยมักทำคือ การรายงานการวิเคราะห์จากผลการสำรวจเฉพาะผลลัพธ์แสดงให้เห็นถึงผลกระทบ แสดงให้เห็นถึงความลำเอียงเป็นพื้นฐานในเรื่อง

ความคลาดเคลื่อนเชิงบวกที่เป็นเท็จ (false positive error) และแสดงให้เห็นถึงความจำเป็นต้องทำการวิจัยเชิงยืนยัน รวมถึงมีความพยายามให้ข้อมูลหรือหลักฐานเชิงโน้มน้าวจากการวิจัยเชิงสำรวจโดยเน้นให้เห็นอัตราการเกิดความคลาดเคลื่อนแบบที่ 1 โดยไม่ขจัดความลำเอียงในการเลือกการรายงานและไม่กำจัดสิ่งที่ต้องทำในการวิจัยเชิงยืนยัน โดยเฉพาะในช่วงคริสต์ทศวรรษที่ผ่านการวิจัยด้านสังคมศาสตร์มีการพัฒนางานวิจัยเชิงสำรวจที่ตอกย้ำถึงความล่อแหลม (unhealthy) ในการวิจัยเชิงสำรวจที่ไม่เน้นหรือไม่ให้ความสำคัญกับการยืนยันผลการสำรวจ รวมถึงมีงานวิจัยบางส่วนที่ยังคงใช้วิธีการทางสถิติที่คลุมเครือระหว่างการสำรวจและการยืนยัน เพราะความคลาดเคลื่อนในการอ้างอิงสามารถเกิดขึ้นกับการทดสอบสมมติฐานทุกวิธี (Kennedy, 2015)

การวิเคราะห์ความเสี่ยง

คำว่า “ความเสี่ยง (risk)” กับ “ความไม่แน่นอน (uncertainty)” มักใช้สลับหรือแทนกันเสมอ แต่ความจริงแล้วไม่เหมือนกัน กล่าวคือ ความเสี่ยงเป็นผลกระทบสะสม (cumulative effect) ของความน่าจะเป็น (probability) ของเหตุการณ์ที่เกิดจากความไม่แน่นอน อาจเป็นผลกระทบเชิงบวกหรือเชิงลบ ส่วนความไม่แน่นอนเป็นเรื่องที่เน้นเฉพาะความน่าจะเป็นของเหตุการณ์ที่ไม่รู้แหล่งที่มาหรือไม่ทราบสาเหตุโดยสิ้นเชิง (Pritchard, 2010: 7)

ความเสี่ยงในแต่ละบริบทมีความแตกต่างกันทั้งด้านความหมาย เป้าหมาย และวิธีในการวัด เช่น ด้านวิศวกรรมให้คำจำกัดความเกี่ยวกับความเสียหาย ด้านเศรษฐศาสตร์ให้คำอธิบายเกี่ยวความไม่มีประโยชน์ (disutility) และด้านการจราจรให้นิยามเกี่ยวกับการเสียชีวิต

ความเสี่ยงในมุมมองดังเดิม หมายถึง สถานการณ์หรือเหตุการณ์ที่อาจเกิดและอาจเกิดขึ้นบ่อย สามารถประเมินได้ด้วยการแจกแจงความน่าจะเป็น (probability distribution) จากเหตุการณ์ที่เกิดขึ้นในอดีต โดยมี 3 องค์ประกอบพื้นฐาน คือ เหตุการณ์ (event) ความน่าจะเป็น (probability) และความรุนแรงหรือผลกระทบหรือผลสืบเนื่อง (severity/impact/consequence) แต่ส่วนใหญ่เน้นไปที่ความเสี่ยงที่เป็นเหตุการณ์ที่ปรากฏให้เห็นจริง (materialize) และต้องเกิดขึ้น (must occur) บางครั้งอาจเป็นเหตุการณ์ที่ไม่ได้วางแผนไว้ (unplanned event) รวมถึงผลสืบเนื่องที่ไม่คาดว่าจะเกิดขึ้น (unexpected consequences) (Pritchard, 2010: 7; Aven, 2011: 16-31; Hopkin, 2017: 15-16)

ความเสี่ยงในแต่ละเหตุการณ์ทำให้เกิดผลกระทบเชิงบวก (positive) และในทางกลับกันก็มีโอกาสที่จะเกิดผลกระทบเชิงลบ (negative) หรือเบี่ยงเบน (deviation) จากสิ่งที่คาดหวังไว้ ทั้งความเสี่ยงและโอกาสมีความเกี่ยวพันกันหรือมีผลกระทบต่อกันอย่างหลีกเลี่ยงไม่ได้ ทำให้ไม่สามารถกำหนดได้ว่าเป็นผลกระทบเชิงบวกหรือเชิงลบของความเสี่ยงหรือโอกาส อีกทั้งมีความเป็นไปได้ที่ความเสี่ยงนั้นเป็นความไม่แน่นอนของผลลัพธ์ (uncertainty of outcome) (Hopkin, 2017: 15-16; Fenton & Neil, 2019: 53)

โดยทั่วไป การวัดความเสี่ยง (risk measure) ภายใต้กรอบของผลกระทบ (impact-based) วัดจากทั้งความน่าจะเป็นและผลกระทบจากมาตรวัด (scale) 1-5 หรือ 1-10 และสามารถเขียนเป็นสูตรอธิบายง่ายๆ ดังนี้ (Fenton & Neil, 2019: 52)

ความเสี่ยง (risk) = ความน่าจะเป็น (probability) X ผลกระทบ (impact)

หรือ

ความเสี่ยง (risk) = ความควรจะเป็น (likelihood) X ความสูญเสีย (loss)

อย่างไรก็ตาม ความเสี่ยงนอกจากจะมีความหมายที่แตกต่างกันแล้ว ยังมีการใช้ดัชนีวัดที่แตกต่างกันด้วย บางนิยามใช้การวัดเพียงมิติเดียว คือ ความไม่แน่นอนของผลลัพธ์หรือการกระทำหรือเหตุการณ์ ซึ่งถูกโจมตีว่าเป็นนิยามที่ใช้การไม่ได้ เพราะไม่สามารถแสดงให้เห็นถึงมิติที่สำคัญ คือ ผลสืบเนื่อง เช่น อัตราการตายจากการจราจรระหว่างปีที่ต่ำกว่า แสดงว่ามีความเสี่ยงต่ำกว่า แต่ในความเป็นจริงมีผู้เสียชีวิตเป็นพันคน ดังนั้น หลายนิยามจึงมีการวัดจากหลายมิติ เช่น เหตุการณ์หรือโครงการที่เป็นจุดเริ่มต้นของผลสืบเนื่องของเหตุการณ์ และความน่าจะเป็นหรือความไม่แน่นอนที่เกิดจากความสัมพันธ์กันระหว่างเหตุการณ์กับผลกระทบ ส่วนค่าที่คาดหมายไว้ว่าจะเกิดความเสี่ยงอาจไม่สัมพันธ์กับเหตุการณ์ที่เกิดขึ้น เช่น บางเหตุการณ์มีความน่าจะเป็นต่ำที่จะเกิด แต่มีผลกระทบสูง เช่น อุบัติเหตุทางนิวเคลียร์

แม้ว่าแนวคิดเกี่ยวกับความเสี่ยงมีความหลากหลายและมีการวัดหรือการประเมินที่แตกต่างกัน แต่สามารถแบ่งได้เป็น 2 กลุ่ม ดังนี้ (Aven, 2011: 16-31)

1. ความเสี่ยงที่เกิดจากความน่าจะเป็นเชิงวัตถุวิสัย (Objective probability) ใช้วิธีการเชิงสถิติเพียงอย่างเดียว (pure statistical approach) เพื่อหาการประมาณการความถูกต้อง (accurately estimate) โดยใช้ช่วงความเชื่อมั่น (confidence interval) จากความสัมพันธ์ของความน่าจะเป็นที่ได้มาจากการอธิบายด้วยความถี่ (relative frequency-interpreted probability) สามารถเขียนเป็นสูตรอธิบายได้ดังนี้

ความเสี่ยง = (A, C, P)

โดยให้ A = เหตุการณ์ C = ผลกระทบ และ P = ความน่าจะเป็น

2. ความเสี่ยงที่เกิดจากความน่าจะเป็นเชิงจิตวิสัย (Subjective probability) หรือความน่าจะเป็นบนฐานของความรู้ (knowledge base probability) โดยใช้วิธีการประเมินความเสี่ยงด้วยการอธิบายความไม่แน่นอน จากจำนวนที่ไม่รู้ในสิ่งที่สนใจ (unknown quantities of interest) สามารถเขียนเป็นสูตรอธิบายได้ดังนี้

ความเสี่ยง = (A, C, U)

โดยให้ A = เหตุการณ์ C = ผลกระทบ และ U = ความไม่แน่นอน

การศึกษาเกี่ยวกับความเสี่ยงยังมีความเห็นที่แตกต่างกันในบางขั้นตอนหรือที่เรียกว่า กระบวนการประเมินความเสี่ยง (risk assessment process) กล่าวคือ สมาชิกของสมาคมวิเคราะห์ความเสี่ยง (Society of Risk Analysis) มองว่า การวิเคราะห์ความเสี่ยง (risk analysis) เป็นคำที่ใช้เรียกแนวคิดทั้งหมดที่ประกอบไปด้วยการประเมินความเสี่ยง (risk assessment) การยอมรับความเสี่ยง (risk perception) การจัดการความเสี่ยง (risk management) การสื่อสารความเสี่ยง (risk communication) และสิ่งที่เกี่ยวข้องกับความเสี่ยงที่กล่าวมาทั้งหมด แต่องค์กรระหว่างประเทศว่าด้วยมาตรฐาน (International Organization for Standardization : ISO) มองว่า การวิเคราะห์ความเสี่ยงรวมถึงการวัดผลความเสี่ยง (risk evaluation) และเป็นส่วนหนึ่งของการประเมินความเสี่ยง แต่ไม่รวมถึงการระบุต้นเหตุของความเสี่ยงให้เป็นส่วนหนึ่งของการวิเคราะห์ความเสี่ยง (Aven, 2011: 4-5, 35)

การวิเคราะห์และวัดผลหรือประเมินความเสี่ยงแบบดังเดิมใช้วิธีการที่เป็นวิทยาศาสตร์แบบเก่า (traditional scientific methods) ภายใต้วัฒนธรรมหรือบริบทของแต่ละศาสตร์ เช่น วิทยาศาสตร์ธรรมชาติ สังคมศาสตร์ คณิตศาสตร์ และทฤษฎีความน่าจะเป็น ที่ต้องได้คำตอบที่มีความเที่ยงตรง (validity) และมีความน่าเชื่อถือ (reliability) ถูกวิพากษ์ว่า ไม่เหมาะที่จะนำมาใช้กับการประเมินหรือทำนายความเสี่ยงที่มีบริบทหรือสถานการณ์ที่มีความหลากหลาย ข้อกำหนดคุณภาพของการประเมินความเสี่ยงที่เป็นวิทยาศาสตร์ต้องมีความเที่ยงตรงและความน่าเชื่อถือที่เหมาะกับแต่ละวัตถุประสงค์หรือสิ่งที่คาดว่าจะเกิดขึ้นในอนาคต (expectations) และการวิเคราะห์หรือประเมินความเสี่ยงด้วยวิธีการแบบเก่ามีข้อบกพร่องที่ทำให้เกิดความผิดพลาดในการอุปนัย (induction) เช่น มีการใช้วิธีการทางสถิติที่แตกต่างกันในการประเมินสถานการณ์ที่เหมือนกัน ค่าความน่าจะเป็น (*p*-value) ไม่สามารถบอกผลกระทบหรือขนาด

ของผลกระทบที่เกิดจากตัวแปรอื่นๆ วิธีการทางสถิติแบบเก่าไม่สามารถทำนายการเกิดวิกฤติทางการเงินหรือการก่อการร้ายที่มีความซับซ้อนมากขึ้น (Aven, 2011: 174-175; Fenton & Neil, 2019: 9-43)

วิธีการวิเคราะห์ความเสี่ยงเกือบทั้งหมดต้องใช้ความเชี่ยวชาญในการตัดสินใจ แม้ว่าจะมีผู้ที่เกี่ยวข้องไม่เห็นด้วยว่าจะเกิดความเสี่ยง แต่ก็ต้องมีการเตรียมการสำหรับการตัดสินใจในขั้นสุดท้าย รวมถึงต้องรับฟังความคิดเห็นจากผู้เชี่ยวชาญจำนวนมาก และต้องยอมรับว่ายากที่วินิจฉัยได้ว่าตัดสินใจได้ดีหรือไม่ดี (Pritchard, 2010: 9, 63) เพราะการนำผลการวิเคราะห์หรือประเมินความเสี่ยงไปใช้ขึ้นอยู่กับปัจจัยด้านการบริหาร ประสบการณ์ ความรู้ มุมมอง และการพิจารณาของผู้ตัดสินใจ ดังนั้น การประเมินความเสี่ยงเป็นศาสตร์หนึ่งด้วยตัวของมันเอง ไม่สามารถหรือไม่น่าจะต้องอิงกรอบประเพณีนิยมของแต่ละศาสตร์ ต้องออกแบบพัฒนาแนวคิด (concepts) หลักการ (principles) ระเบียบวิธีการ (methods) และตัวแบบ (models) ในการกำหนดลักษณะ (nature) และขอบเขต (extent) ของความเสี่ยงภายใต้แต่ละบริบทของการตัดสินใจ โดยมีขั้นตอนหลักๆ ดังนี้ (Aven, 2011: 174-177)

1. การระบุหรือกำหนดต้นเหตุ (Source Identification) ของความเสี่ยง ได้แก่ ความอันตราย (hazard) ภัยคุกคาม (threat) โอกาส (opportunity)

2. การวิเคราะห์ (Analysis) เพื่อหาสาเหตุ (cause) ผลสืบเนื่อง (consequence) รวมถึงช่องโหว่หรือจุดเปราะบาง (vulnerability)

3. การอธิบายความเสี่ยง (Risk Description) โดยใช้ความน่าจะเป็น (probability) และค่าที่คาดหวัง (expected value) บรรยายหรือวาดภาพแสดงให้เห็นความเสี่ยง (risk picture)

4. กำหนดปัจจัยที่เป็นความไม่แน่นอน (Uncertainty Factors) สำหรับการประเมิน

5. วัดผลความเสี่ยง (Risk Evaluations) โดยการเปรียบเทียบกับเกณฑ์ (criteria) ของระดับความเสี่ยงที่ยอมรับได้ (risk tolerability)

การวิเคราะห์ความเสี่ยงมีความแตกต่างกันในแต่ละความหมายและบริบท มีการพัฒนาวิธีการและดัชนีในการวิเคราะห์ไม่เหมือนกัน มีทั้งวิธีการที่เป็นตัวแบบเชิงสถิติ และวิธีการอธิบายด้วยความรู้ โดยไม่ใช้ตัวแบบเชิงสถิติ ส่วนการวิเคราะห์ความเสี่ยงด้วยตัวแบบเชิงสถิติมีทั้งใช้สถิติเชิงพรรณนาและสถิติเชิง

อ้างอิง การใช้สถิติเชิงอ้างอิงมีทั้งสถิติแบบความถี่ (frequentist statistics) และสถิติแบบเบย์ (bayesian statistics)

สรุป ความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยเชิงปริมาณที่ทำให้เกิดความไม่ถูกต้องในการปฏิเสธหรือยอมรับสมมติฐานมีลักษณะเป็นห่วงโซ่ต่อเนื่องกันและเกิดมาจากหลายสาเหตุ เช่น ระดับของการวัด ประชากร ขนาดตัวอย่าง การเลือกตัวอย่าง ข้อมูล การคำนวณ การเลือกวิธีการทดสอบวิธีการทางสถิติ เลือกใช้เครื่องมือวัดที่ไม่เหมาะสม และการยกเว้นการใช้กฎต่างๆ แต่ที่พบมากที่สุด คือ ความคลาดเคลื่อนที่เกิดจากเครื่องมือ (instrument error) และความคลาดเคลื่อนที่เกิดการเลือกตัวอย่าง (sampling error) ความคลาดเคลื่อนที่เกิดขึ้นไม่สามารถกำจัดออกไปได้จากการวิจัย แต่สามารถลดผลกระทบที่เกิดขึ้นได้ เช่น การสร้างแบบสอบถามให้มีคุณภาพและชัดเจน และตรวจสอบข้อมูล ให้ถูกต้องและสมบูรณ์ (Henry, 1990; Biemer & others, 1991; Bevington & Robinson, 1992; Thye, 2000; Holmes, 2004; Groves cited by Scheaffer, Mendenhall III, & Ott, 2006; Buonaccorsi, 2010; Berendsen,2011; Thompson, 2012; Babbie, 2013)

โอกาสของการเกิดความคลาดเคลื่อนสามารถเกิดขึ้นได้จากทุกขั้นตอน หากจำแนกสาเหตุจากต้นกำเนิดสามารถแบ่งได้เป็น 2 แหล่ง คือ เครื่องมือวัด (instrument) และประชากร (population) งานวิจัยบางเรื่องอาจเกิดตั้งแต่เริ่มต้นในขั้นตอนของการสร้างเครื่องมือวัด ไม่มีการตรวจสอบคุณภาพทั้งด้านความเที่ยงตรงและความน่าเชื่อถือ และการเลือกประชากรที่ใช้ในการศึกษา โดยไม่ทราบจำนวนประชากรหรือขอบเขตของประชากรจึงไม่สามารถใช้สถิติแบบมีพารามิเตอร์ที่ใช้อ้างอิงถึงประชากรที่ใช้การศึกษางานวิจัยบางเรื่องอาจเกิดในขั้นตอนการเลือกใช้สถิติไม่ถูกต้อง โดยเฉพาะงานวิจัยที่เลือกใช้สถิติแบบมีพารามิเตอร์ เพราะข้อมูลที่ใช้ในการวิเคราะห์ไม่เป็นไปตามข้อตกลงหรือเงื่อนไขในการใช้สถิติ (violate assumption) อาจทำให้เกิดความคลาดเคลื่อนแบบ β โดยยอมรับสมมติฐานหลักที่เป็นเท็จโดยปฏิเสธสมมติฐานทางเลือกที่เป็นจริง (type II error)

มีผลการศึกษาจำนวนมากที่ตรวจสอบพบว่า รายงานผลการศึกษาที่ใช้ข้อมูลจากการสำรวจมีความคลาดเคลื่อน และการวัดที่ขาดความน่าเชื่อถือเป็นสาเหตุทำให้ได้ค่าประมาณการเบี่ยงเบนไปจากความเป็นจริง และนำไปสู่การเกิดความคลาดเคลื่อนของผลการวิจัยตามมา เครื่องมือวัดจึงมีอิทธิพลต่อข้อมูล เพื่อให้

ได้ข้อมูลที่ตรงกับความเป็นจริง เครื่องมือวัดต้องมีความเที่ยงตรง (validity) และความน่าเชื่อถือ (reliability) หากเครื่องมือวัดมีความคลาดเคลื่อน ไม่สามารถวัดได้อย่างเที่ยงตรงก็จะทำให้เกิดความคลาดเคลื่อนแบบเป็นระบบ (systematic error) และไม่สามารถวัดได้อย่างน่าเชื่อถือก็จะทำให้เกิดความคลาดเคลื่อนแบบสุ่ม (random error) โดยในกระบวนการพัฒนาและการตรวจสอบเครื่องมือเพื่อลดความคลาดเคลื่อนจากการวัด มี 2 ดัชนีที่แสดงถึงความมีคุณภาพของเครื่องมือ คือ ความเที่ยงตรง และความน่าเชื่อถือ แม้ว่าเครื่องมือมีความเที่ยงตรงและมีความน่าเชื่อถือแล้ว ก่อนทำการรวบรวมข้อมูลก็ต้องนำเอาเครื่องมือวิจัยไปทำการทดสอบก่อนอีกครั้ง เพื่อตรวจสอบปัญหาเกี่ยวกับความเข้าใจของแต่ละคำถามจากผู้ตอบคำถาม ในพื้นที่วิจัยจริงกับกลุ่มคนที่มีลักษณะที่เหมือนกับประชากรที่ทำการศึกษา (Bollinger, 1998; Groves cited by Scheaffer, Mendenhall III, & Ott, 2006; Kimberlin & Winterstein, 2008; Osborne, 2008; Lash, Fox, & Fink, 2009; Drost, 2011; Schennach, 2012 ; Thompson, 2012; Kumar, 2014; Taherdoost, 2016)

ความเที่ยงตรงและความน่าเชื่อถือในการวัดจะมีความสัมพันธ์กับขนาดกลุ่มตัวอย่างที่มีขนาดใหญ่กว่า เนื่องจากจะช่วยแก้ไขปัญหาความแปรปรวนที่เพิ่มขึ้นได้หรือทำให้ความแปรปรวนในการสุ่มลดลง แม้ว่ากลุ่มตัวอย่างมีขนาดใหญ่ แต่หากเครื่องมือวัดที่นำมาใช้ในการเก็บรวบรวมข้อมูลมีความน่าเชื่อถือน้อย ข้อมูลที่ได้จากการสังเกตก็จะมีความแปรปรวนมากขึ้น เมื่อข้อมูลมีความแปรปรวนมากขึ้นก็จะมีผลทำให้เกิดการปฏิเสธสมมติฐานหลักยากมากขึ้นตามมาแม้ว่าในความเป็นจริงจะมีความสัมพันธ์เกิดขึ้นจริง ดังนั้นในการออกแบบเลือกตัวอย่างจึงต้องเลือกใช้วิธีการที่เหมาะสม และได้คำตอบจากตัวแทนของประชากรอย่างแท้จริง (Henry, 1990; Biemer & others, 1991)

วิธีการเลือกตัวอย่างหรือการออกแบบเลือกตัวอย่าง (sampling design) เป็นกระบวนการที่ให้ได้มาแต่ละตัวอย่าง (*s*) ด้วยความน่าจะเป็น P(*s*) การลดความคลาดเคลื่อนหรือความลำเอียงในการประมาณจึงต้องมีความรอบครอบในการออกแบบเลือกตัวอย่างดังนี้ (Goldstein, 1989; Flynn, Sakakibara, Schroeder, Bates, & Flynn, 1990; Scheaffer, Mendenhall III, & Ott, 2006; Burdick & Watrin, 2008; Burmeister & Aitken, 2012; Thompson, 2012)

1. ประชากร (population) การเลือกตัวอย่างมีวัตถุประสงค์เพื่อใช้ในการอ้างอิง (inference) จากตัวอย่างไปสู่ประชากรที่ทราบจำนวน (finite population) ด้วยวิธีการประมาณ (estimate) จากค่าพารามิเตอร์ของประชากร (population parameters) หากกระบวนการการออกแบบและการดำเนินการในการเลือกตัวอย่างไม่ถูกต้อง นอกจากจะมีผลต่อความเที่ยงตรงภายนอก (external validity) หรือความสามารถในการสร้างข้อสรุป (generalizability) ไปสู่ประชากรเป้าหมายแล้ว ยังมีผลต่อการนำเอาผลการศึกษาไปใช้สร้างความสัมพันธ์เชิงสาเหตุในสถานการณ์อื่นๆ หรือประชาชนกลุ่มอื่นๆ รวมถึงมีผลกระทบโดยตรงต่อความถูกต้องในการสรุปเชิงสถิติอีกด้วย

2. ขนาดตัวอย่าง (sample size) มีผลอย่างมีนัยสำคัญต่อความเที่ยงตรง ความเสี่ยงของการเกิดความคลาดเคลื่อน และข้อค้นพบที่ได้จากการวิจัย การทดสอบสมมติฐาน การยืนยัน หรือตรวจสอบทฤษฎี นักวิจัยอาจต้องตัดสินใจเลือกวิธีคำนวณขนาดตัวอย่างด้วยวิธีการใดวิธีการหนึ่งระหว่างการทดสอบนัยสำคัญทางสถิติและทดสอบอำนาจจำแนกทางสถิติ

3. วิธีการเลือกตัวอย่าง (sampling) การเลือกตัวอย่างแบบสุ่มจะช่วยป้องกันไม่ให้เกิดความลำเอียง ส่วนตัวอย่างที่มาจากการสุ่มแบบตามสะดวกจะมีความลำเอียงสูงมาก

สถิติเชิงอ้างอิงแบบมีพารามิเตอร์ (parametric statistics) คือ สถิติที่มีค่าตัวเลขทางประชากรที่ใช้ในการอ้างอิงไปสู่คุณลักษณะของประชากร ความคลาดเคลื่อนเกิดมาจากข้อมูลหรือคำตอบที่ได้มาจากการเลือกตัวอย่างแบบสุ่มทั้งจากประชากรที่ทราบจำนวน (finite population) และประชากรที่ไม่ทราบจำนวน (infinite population) การเลือกตัวอย่างมีวัตถุประสงค์เพื่อใช้ในการอ้างอิง (inference) จากตัวอย่าง (sample) ไปสู่ประชากรด้วยวิธีการประมาณ (estimate) จากค่าพารามิเตอร์ของประชากร (population parameters) (Biemer & others, 1991; Scheaffer, Mendenhall III, & Ott, 2006; Healey, 1999; Sheskin, 2007; Field, 2009; Yang, 2010; Evans, 2014)

สถิติแบบมีพารามิเตอร์ส่วนใหญ่นอกจากวางอยู่บนฐานของข้อมูลที่มีระดับการวัดตั้งแต่ช่วงมาตราขึ้นไปและข้อมูลต้องมีการแจกแจงแบบปกติจากบัญชีรายชื่อขนาดใหญ่ (large catalogue) แล้ว สถิติแต่ละประเภทยังมีข้อตกลงเฉพาะเพิ่มเติมอีก เช่น การวิเคราะห์เปรียบเทียบค่าเฉลี่ยระหว่างกลุ่มตัวอย่าง ข้อมูลแต่ละกลุ่มตัวอย่างต้องมาจากประชากรที่มีความเป็นแบบเดียวกันของความแปรปรวน (homogeneity of variance) การวิเคราะห์ค่าเฉลี่ยแบบทดสอบซ้ำ ข้อมูลของแต่ละตัวอย่างต้องมีความ

เป็นอิสระจากกัน (independence) และการวิเคราะห์ตัวแบบการทำนาย ข้อมูลระหว่างตัวแปรที่ใช้ทำนายและตัวแปรที่ถูกทำนายต้องมีความสัมพันธ์กันเชิงเส้นตรง (linearity) หากไม่ให้ความสนใจข้อตกลงเหล่านี้ผลการวิเคราะห์จะไม่มีความถูกต้องและทำให้ได้ข้อสรุปที่ผิด แต่ยังมีนักวิจัยส่วนใหญ่ไม่ให้ความสนใจข้อตกลงของสถิติแบบมีพารามิเตอร์ รวมถึงมีความเห็นและเชื่อกันมาอย่างยาวนานว่า ไม่มีเหตุผลอื่นที่จะทำให้เกิดการละเมิด ข้อตกลงเกี่ยวกับการทดสอบสถิติแบบมีพารามิเตอร์ หากข้อมูลที่ใช้มีระดับการวัดเป็นช่วงมาตราหรือสัดส่วนมาตรา (Healey, 1999; Field, 2009; Yang, 2010; Evans, 2014)

แม้ว่าสถิติแบบไม่มีพารามิเตอร์จะมีข้อตกลงเล็กน้อยหรือไม่มีข้อตกลงเลย (free assumption) ก็ไม่ได้หมายความว่าจะไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการสรุปผลการการวิจัย หากการออกแบบการวิจัยทั้งด้านเครื่องมือวัด ประชากร และข้อมูล ตรงกับข้อตกลง (met assumptions) ในการใช้สถิติแบบมีพารามิเตอร์ได้ แต่เลือกใช้การทดสอบด้วยสถิติแบบไม่มีพารามิเตอร์จะทำให้อำนาจการทดสอบต่ำกว่าการทดสอบด้วยสถิติแบบมีพารามิเตอร์ อาจทำให้เกิดความคลาดเคลื่อนแบบ α โดยปฏิเสธสมมติฐานหลักที่เป็นจริงและยอมรับสมมติฐานทางเลือกที่เป็นเท็จ (type I error) ซึ่งจะนำไปสู่การสรุปผลการวิจัยที่ไม่ถูกต้องตามมาได้เช่นกัน

การละเมิดข้อตกลงบางครั้งอาจไม่มีผลต่อข้อสรุปของการวิจัยแต่อย่างไร แต่บางครั้งอาจมีผลต่อการแปลความหมายของผลการวิจัยเป็นอย่างมาก และไม่น่าที่จะทำให้ได้ผลการวิจัยที่ถูกต้องและได้ข้อสรุปที่น่าเชื่อถือ เกิดความคลาดเคลื่อนแบบที่ 1 และความคลาดเคลื่อนแบบที่ 2 หรือทำให้การประมาณการค่านัยสำคัญหรือขนาดอิทธิพลต่ำกว่าหรือมากกว่า แต่เรากลับยัดเยียดความถูกต้องไว้ในผลลัพธ์ ข้อสรุป และการยืนยันผลของการวิจัย โดยไม่ได้ทำการทดสอบตามหลักสถิติว่า เป็นไปตามข้อตกลงเบื้องต้นหรือไม่ (Osborne & Waters, 2002; Sheskin, 2007; Ghasemi & Zahediasl, 2012)

การวัดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย

จากการทบทวนวรรณกรรม แนวคิด และทฤษฎีที่เกี่ยวข้อง พบว่า ความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยเชิงปริมาณเกิดมาจากความคลาดเคลื่อนมีผลมาจากการเลือกตัวอย่าง (sampling errors) และความคลาดเคลื่อนที่ไม่ได้มีผลมาจากการเลือกตัวอย่าง (nonsampling error) ซึ่งเป็นความลำเอียงและละเมิดข้อตกลง ทำให้การทดสอบนัยสำคัญทางสถิติ (statistical significance test) และการทดสอบอำนาจจำแนกทางสถิติ (statistical power test) ด้วยสถิติแบบมีพารามิเตอร์ในการทดสอบสมมติฐาน เกิดความคลาดเคลื่อนแบบที่ 1 และความคลาดเคลื่อนแบบที่ 2 และทำให้การแปลความหมายของผลการวิจัยไม่ถูกต้อง ดังนั้นในการศึกษาครั้งนี้จึงนำเอาปัจจัยที่ใช้ในการทดสอบความเสี่ยงของการเกิดความคลาดเคลื่อนมาทำการวัดดังต่อไปนี้

ตารางที่ 2.1 แสดงการวัดตัวแปรความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย

มิติของการวัด	แนวคิด	การวัด	งานที่สนับสนุน
เครื่องมือวัด	หากเครื่องมือวัดมีความคลาดเคลื่อนไม่สามารถวัดได้อย่างเที่ยงตรงก็จะทำให้เกิดความคลาดเคลื่อนแบบเป็นระบบและไม่สามารถวัดได้อย่างน่าเชื่อถือก็จะทำให้เกิดความคลาดเคลื่อนแบบสุ่ม	เครื่องมือวัดมีการตรวจสอบความเที่ยงตรงและความน่าเชื่อถือหรือไม่มีการตรวจสอบความเที่ยงตรงและความน่าเชื่อถือ	Kimberlin & Winterstein, 2008; Taherdoost, 2016
ระดับการวัด	สถิติแบบมีพารามิเตอร์ส่วนใหญ่วางอยู่บนฐานของข้อมูลแบบค่าต่อเนื่องหรือเป็นแบบเมตริกที่มีระดับการวัดตั้งแต่ช่วงมาตราขึ้นไป	ข้อมูลที่ใช้ในการวิเคราะห์ด้วยสถิติแบบมีพารามิเตอร์เป็นข้อมูลแบบค่าต่อเนื่องหรือเป็นข้อมูลแบบค่าแบ่งกลุ่ม	Healey, 1999; Sheskin, 2007; Field, 2009; Yang, 2010; Evans, 2014; Verma, & Abdel-Salam, 2019;
ประชากร	เป้าหมายที่สำคัญของการใช้สถิติ คือ การนำเอาข้อมูลหรือสารสนเทศจากตัวอย่างอ้างอิงไปสู่การอธิบายหรือสรุปประชากร และตัวอย่างที่สุ่มมาจาก	ประชากรที่ใช้ในการวิจัยเป็นประชากรที่ทราบจำนวนหรือเป็น	Bross, 1954; Holmes, 2004; Scheaffer, Mendenhall III, &

มิติของการวัด	แนวคิด	การวัด	งานที่สนับสนุน
	ประชากรที่ไม่ทราบจำนวนทำให้ได้ข้อมูลที่มีความคลาดเคลื่อน	ประชากรที่ไม่ทราบจำนวน	Ott, 2006; Burdick & Watrin, 2008
การคำนวณขนาดตัวอย่าง	การทดสอบสมมติฐานที่เน้นการยืนยัน หรือตรวจสอบทฤษฎี นักวิจัยอาจต้องตัดสินใจเลือกวิธีคำนวณขนาดตัวอย่าง ด้วยวิธีการใดวิธีการหนึ่งระหว่างการทดสอบนัยสำคัญทางสถิติและทดสอบอำนาจจำแนกทางสถิติ หากต้องการป้องกันการเกิดความคลาดเคลื่อนแบบที่ 1 จึงต้องเลือกกำหนดค่านัยสำคัญให้ต่ำ ซึ่งจะมีผลทำให้ต้องใช้ขนาดตัวอย่างเพิ่มขึ้น แต่ขณะเดียวกันจะทำให้เพิ่มโอกาสการเกิดความคลาดเคลื่อนแบบที่ 2 มากขึ้น และหากต้องการป้องกันการเกิดความคลาดเคลื่อนแบบที่ 2 ต้องเลือกกำหนดค่าอำนาจจำแนกให้ต่ำ ซึ่งมีผลทำให้ต้องใช้ขนาดตัวอย่างเพิ่มขึ้นเช่นเดียวกัน การกำหนดขนาดตัวอย่างด้วยวิธีการทางสถิติ จะทำให้ใช้จำนวนตัวอย่างไม่มากและไม่น้อยเกินไป ทำให้เกิดความคลาดเคลื่อนน้อยลง มีความเชื่อมั่นหรือความแม่นยำมากขึ้น	การคำนวณขนาดตัวอย่างใช้วิธีการทางสถิติหรือไม่ใช้วิธีการทางสถิติ	Goldstein, 1989; Holmes, 2004; Youssef, 2011; Burmeister & Aitken, 2012
การเลือกตัวอย่าง	วิธีการเลือกตัวอย่างหรือการออกแบบเลือกตัวอย่างเป็นกระบวนการที่ทำให้ได้มาแต่ละตัวอย่างด้วยความน่าจะเป็น การเลือกตัวอย่างแบบใช้หลักความน่าจะเป็นหรือแบบสุ่มจะช่วยป้องกันไม่ให้เกิดความลำเอียง ส่วนตัวอย่างแบบไม่ใช้หลักความน่าจะเป็นที่มาจากการสุ่มแบบตามสะดวกจะมีความลำเอียงสูงมาก	การเลือกตัวอย่างแบบใช้หลักความน่าจะเป็นหรือไม่ใช้หลักความน่าจะเป็น	Flynn, Sakakibara, Schroeder, Bates, & Flynn, 1990; Thompson, 2012; Blair & Blair, 2015
การตรวจสอบข้อมูล	นักวิจัยส่วนใหญ่ไม่ให้ความสนใจข้อตกลงของสถิติแบบมีพารามิเตอร์ ดังนั้นผลการวิเคราะห์จึงไม่มีความถูกต้อง การตรวจสอบข้อมูลให้ตรงกับข้อตกลงของสถิติทดสอบแต่ละประเภทจึงมีความสำคัญทั้งด้านความถูกต้องตามข้อเท็จจริงของข้อมูล และเป็น	มีการตรวจสอบข้อมูลตามข้อตกลงของสถิติแต่ละประเภทหรือไม่มีการตรวจสอบตาม	Allison, 2002; Hair & Others, 2006; Field, 2009; Garson, 2012; Verma, &

มิติของการวัด	แนวคิด	การวัด	งานที่สนับสนุน
	พื้นฐานของการตัดสินใจเลือกใช้สถิติเชิงอ้างอิงให้ถูกต้องตามหลักสถิติ เพื่อให้เกิดความน่าเชื่อถือในการสรุปและอ้างอิงไปถึงประชากร	ข้อตกลงของสถิติแต่ละประเภท	Abdel-Salam, 2019
สถิติ	โอกาสของการเกิดความคลาดเคลื่อนสามารถเกิดขึ้นได้จากทุกขั้นตอน โดยเฉพาะงานวิจัยที่เลือกใช้สถิติแบบมีพารามิเตอร์ เพราะข้อมูลที่ใช้ในการวิเคราะห์ไม่เป็นไปตามข้อตกลงหรือเงื่อนไขในการใช้สถิติ อาจทำให้เกิดความคลาดเคลื่อนแบบที่ 2 โดยยอมรับสมมติฐานหลักที่เป็นเท็จโดยปฏิเสธสมมติฐานทางเลือกที่เป็นจริง	ใช้สถิติแบบมีพารามิเตอร์หรือสถิติแบบไม่มีพารามิเตอร์ในการวิเคราะห์ข้อมูล	Verma, & Abdel-Salam, 2019
สมมติฐาน	ผลการวิจัยไม่เคยมีความแน่นอน ต้องยอมรับว่าการทดสอบสมมติฐานด้วยวิธีการทางสถิติมีความไม่แน่นอน เพราะเป็นการค้นพบจากความน่าจะเป็นทางสถิติที่ได้จากการสิ่งที่สังเกต ความคลาดเคลื่อนแบบที่ 1 หรือความคลาดเคลื่อนที่เกิดจากความน่าจะเป็นของความไม่ถูกต้องในการปฏิเสธสมมติฐานหลักเกิดได้จากหลายสาเหตุ	เป็นงานวิจัยที่มีการทดสอบสมมติฐานหรือไม่มีการทดสอบสมมติฐาน	Bross, 1954; Holmes, 2004
ความเสี่ยงของการเกิดความคลาดเคลื่อน	ความเสี่ยงเป็นความไม่แน่นอนของผลลัพธ์ การประเมินความเสี่ยงเป็นศาสตร์หนึ่งด้วยตัวของมันเอง ไม่น่าจะต้องอิงกรอบประเพณีนิยมของแต่ละศาสตร์ ต้องออกแบบพัฒนาแนวคิด หลักการ ระเบียบวิธีการ และตัวแบบความเสี่ยงภายใต้แต่ละบริบทของการตัดสินใจ เป้าหมายที่สำคัญของการใช้สถิติ คือ การนำเอาสารสนเทศจากตัวอย่างอ้างอิงไปสู่การอธิบายหรือสรุปประชากร การละเมิดข้อตกลงบางครั้งอาจไม่มีผลต่อข้อสรุปของการวิจัย แต่อย่างไร แต่บางครั้งอาจมีผลต่อการแปลความหมายของผลการวิจัยเป็นอย่างมาก และไม่น่าที่จะทำให้ได้ผลการวิจัยที่ถูกต้องและได้ข้อสรุปที่น่าเชื่อถือ เกิดความคลาดเคลื่อนแบบที่ 1 และความคลาดเคลื่อนแบบที่ 2	งานวิจัยมีการละเมิดข้อตกลงที่ทำให้เกิดความเสี่ยงหรือไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อนของผลการวิจัย	Bross, 1954; Bollinger, 1998; Osborne & Waters, 2002; Scheaffer, Mendenhall III, & Ott, 2006; Osborne, 2008; Aven, 2011; Ghasemi & Zahediasl, 2012; Schennach, 2012; Hopkin, 2017; Fenton & Neil, 2019;

กรอบแนวคิดการวิจัย

การวิจัยครั้งนี้นำเอาปัจจัยที่มีผลต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมาสร้างเป็นตัวแบบเพื่อวิเคราะห์หาความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย โดยแบ่งองค์ประกอบที่เป็นสาเหตุทำให้เกิดความเสี่ยงของการเกิดความคลาดเคลื่อนดังต่อไปนี้

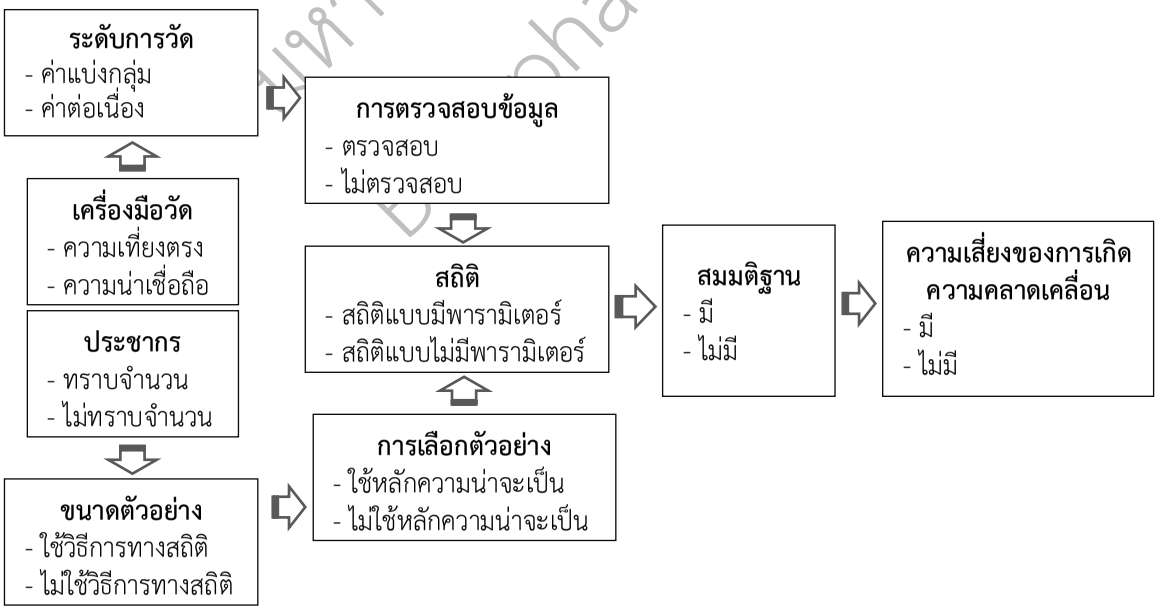
1. เครื่องมือวัด เพื่อให้ได้ข้อมูลที่ตรงกับความเป็นจริง เครื่องมือวัดต้องมีความเที่ยงตรง (validity) และความน่าเชื่อถือ (reliability) หากเครื่องมือวัดมีความคลาดเคลื่อนไม่สามารถวัดได้อย่างเที่ยงตรงก็จะทำให้เกิดความคลาดเคลื่อนแบบเป็นระบบและไม่สามารถวัดได้อย่างน่าเชื่อถือก็จะทำให้เกิดความคลาดเคลื่อนแบบสุ่ม การวิเคราะห์ข้อมูลด้วยสถิติเพื่อทดสอบสมมติฐานการวิจัย การตัดสินใจเลือกใช้สถิติแต่ละประเภทมีข้อตกลงเบื้องต้น (basic assumptions) เกี่ยวกับข้อมูลที่ได้จากการวัด กล่าวคือ สถิติแบบมีพารามิเตอร์ ข้อมูลควรเป็นค่าต่อเนื่อง (continuous data) หรือเป็นข้อมูลแบบเมตริกที่มีระดับการวัดตั้งแต่ช่วงมาตราขึ้นไป และมีการตรวจสอบข้อมูล (data exploration) ตามข้อตกลงของสถิติแต่ละประเภท เช่น การแจกแจงปกติ ค่าขาดหาย ความแปรปรวน ความเป็นเส้นตรง และความสุ่ม หากไม่ตรงกับข้อตกลงต้องเลือกใช้สถิติแบบไม่มีพารามิเตอร์ เพราะการละเมิดข้อตกลงบางครั้งอาจไม่มีผลต่อข้อสรุปของการวิจัยแต่อย่างไร แต่บางครั้งอาจมีผลต่อการแปลความหมายของผลการวิจัยเป็นอย่างมาก และไม่น่าที่จะทำให้ได้ผลการวิจัยที่ถูกต้องและได้ข้อสรุปที่น่าเชื่อถือ เกิดความคลาดเคลื่อนแบบที่ 1 และความคลาดเคลื่อนแบบที่ 2

2. ประชากร การวิจัยทางสังคมศาสตร์ส่วนใหญ่ใช้ประชากรเป็นตัวอย่างในการวิจัย (subject) การเลือกตัวอย่าง (sampling) มีวัตถุประสงค์เพื่อใช้ในการอ้างอิงจากตัวอย่างไปสู่ประชากรที่ทราบจำนวน (finite population) ด้วยวิธีการประมาณจากค่าพารามิเตอร์ของประชากร การกำหนดขนาดตัวอย่าง (sample size determination) มีผลอย่างมีนัยสำคัญต่อความเที่ยงตรงและความเสี่ยงของการเกิดความคลาดเคลื่อนในการทดสอบสมมติฐาน นักวิจัยควรใช้คำนวณขนาดตัวอย่างด้วยวิธีการทางสถิติ (statistically determined sample size) เพราะหากต้องการป้องกันการเกิดความคลาดเคลื่อนแบบที่ 1 จึงต้องเลือกกำหนดค่านัยสำคัญให้ต่ำ ซึ่งจะมีผลทำให้ต้องใช้ขนาดตัวอย่างเพิ่มขึ้น แต่ขณะเดียวกันจะทำให้

เพิ่มโอกาสการเกิดความคลาดเคลื่อนแบบที่ 2 มากขึ้น และหากต้องการป้องกันการเกิดความคลาดเคลื่อนแบบที่ 2 ต้องเลือกกำหนดค่าอำนาจจำแนกให้ต่ำ ซึ่งมีผลทำให้ต้องใช้ขนาดตัวอย่างเพิ่มขึ้นเช่นเดียวกัน และการเลือกตัวอย่างแบบใช้หลักความน่าจะเป็น (probability sampling) จะช่วยป้องกันไม่ให้เกิดความลำเอียง ส่วนการเลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น (nonprobability sampling) จะมีความลำเอียงสูงมาก

3. สถิติ การเลือกใช้สถิติแต่ละประเภทในการทดสอบสมมติฐานมีความสัมพันธ์กับข้อมูลและการเลือกตัวอย่าง กล่าวคือ หากการออกแบบการวิจัยทั้งด้านเครื่องมือวัด ข้อมูล ประชากร และกลุ่มตัวอย่างตรงกับข้อตกลงในการใช้สถิติแบบมีพารามิเตอร์ แต่เลือกใช้สถิติแบบไม่มีพารามิเตอร์จะทำให้ อำนาจการทดสอบต่ำกว่าการทดสอบด้วยสถิติแบบมีพารามิเตอร์ อาจทำให้เกิดความคลาดเคลื่อนแบบ α โดยปฏิเสธสมมติฐานหลักที่เป็นจริงและยอมรับสมมติฐานทางเลือกที่เป็นเท็จ (type I error) และหากเลือกใช้สถิติแบบมีพารามิเตอร์ แต่ข้อมูลที่ใช้ในการวิเคราะห์ไม่เป็นไปตามข้อตกลงในการใช้สถิติอาจทำให้เกิดความคลาดเคลื่อนแบบ β โดยยอมรับสมมติฐานหลักที่เป็นเท็จโดยปฏิเสธสมมติฐานทางเลือกที่เป็นจริง (type II error)

จากแนวคิดที่กล่าวมา นำเอาองค์ประกอบที่เป็นสาเหตุทำให้เกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมาสร้างเป็นกรอบแนวคิดในการวิจัยได้ดังภาพที่ 2.1



ภาพที่ 2.2 แสดงกรอบแนวคิดในการวิจัย

นิยามศัพท์เชิงปฏิบัติการ

1. ความเสี่ยง หมายถึง โอกาสที่ผลการวิเคราะห์หรือทดสอบสมมติฐานงานวิจัยเชิงปริมาณเกิดคลาดเคลื่อนแบบใดแบบหนึ่ง และนำไปสู่การสรุปผลการวิจัยที่ไม่ถูกต้อง จากการละเมิดข้อตกลง (violate assumption) ของการออกแบบการวิจัยตามระเบียบวิธีการวิจัยเชิงปริมาณ

2. ความคลาดเคลื่อน หมายถึง การยอมรับสมมติฐานทางเลือกที่เป็นเท็จโดยปฏิเสธสมมติฐานหลักที่เป็นจริง หรือที่เรียกว่า ความคลาดเคลื่อนแบบที่ 1 (type I error) และการยอมรับสมมติฐานหลักที่เป็นเท็จโดยปฏิเสธสมมติฐานทางเลือกที่เป็นจริง หรือที่เรียกว่า ความคลาดเคลื่อนแบบที่ 2 (type II error)

3. งานวิจัยเชิงปริมาณ หมายถึง วิทยานิพนธ์หรือดุษฎีนิพนธ์ระดับปริญญาโทหรือปริญญาเอกที่ใช้วิธีการวิเคราะห์ผลการศึกษาด้วยสถิติเชิงพรรณนาและสถิติเชิงอ้างอิง

4. สาขาสังคมวิทยา หมายถึง การจัดหมวดหมู่สาขาวิชาของวิทยานิพนธ์หรือดุษฎีนิพนธ์ในระบบฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทย (ThaiLIS-Thai Library Integrated System) ของสำนักงานคณะกรรมการการอุดมศึกษา

5. เครื่องมือวัด หมายถึง การสร้างแบบสอบถามเพื่อใช้การรวบรวมข้อมูล โดยมีการตรวจสอบความเที่ยงตรง (validity) และความน่าเชื่อถือได้ (reliability)

6. ระดับการวัด หมายถึง การเก็บตามรายการตัวแปรเป็นข้อมูลแบบค่าต่อเนื่อง (continuous data) หรือค่าไม่ต่อเนื่อง (discontinuous data) ดังนี้

   6.1 ข้อมูลแบบค่าต่อเนื่อง ได้แก่ การวัดแบบช่วงมาตรา และการวัดแบบสัดส่วนมาตรา

   6.2 ข้อมูลแบบค่าไม่ต่อเนื่อง ได้แก่ การวัดแบบนามมาตรา และการวัดแบบอันดับมาตรา

7. การตรวจสอบข้อมูล หมายถึง การตรวจสอบหรือสำรวจข้อมูล (data exploration) ตามข้อตกลงเบื้องต้น (basic assumption) ของสถิติแต่ละประเภทที่ใช้ในการวิเคราะห์ข้อมูล ได้แก่ การตรวจสอบค่า

สูญหาย (missing analysis) การตรวจสอบการแจกแจงข้อมูลปกติ (normal distribution) การตรวจสอบความเป็นแบบเดียวกันของความแปรปรวน (homogeneity of variance) การตรวจสอบความเป็นแบบเดียวกันของความแปรปรวนของความคลาดเคลื่อน (homoscedasticity) การตรวจสอบความเป็นเส้นตรง (linearity) การตรวจสอบภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรอิสระหลายตัวแปร (multicollinearity) การตรวจสอบความสัมพันธ์ภายในตัวแปร (autocorrelation) การตรวจสอบความสุ่ม (randomness) และการตรวจสอบค่าส่วนเหลือ (residuals)

8. ประชากร หมายถึง จำนวนประชากรเป้าหมายเป็นประชากรที่ทราบจำนวน (finite population) หรือไม่ทราบจำนวน (infinite population)

9. ขนาดตัวอย่าง หมายถึง การคำนวณขนาดตัวอย่างแบบใช้วิธีการทางสถิติ (statistically determined sample size) หรือไม่ใช้วิธีการทางสถิติ (nonstatistically determined sample size) ดังนี้

9.1 การคำนวณขนาดตัวอย่างแบบใช้วิธีการทางสถิติ คือ วิธีการคำนวณขนาดตัวอย่างที่มีการกำหนดช่วงความเชื่อมั่น (confidence interval) หรือกำหนดขนาดผลกระทบ (effect size) แบบใช้ตารางสำเร็จรูปหรือใช้สูตรคำนวณ หรือใช้โปรแกรมคอมพิวเตอร์

9.2 การคำนวณขนาดตัวอย่างแบบไม่ใช้วิธีการทางสถิติ คือ วิธีการคำนวณขนาดตัวอย่างที่ไม่มีการกำหนดความเชื่อมั่น หรือกำหนดขนาดผลกระทบ ขึ้นอยู่กับการตัดสินใจของนักวิจัย ได้แก่ การสำมะโน (census) การเลียนแบบ (imitating) การกำหนดแบบคร่าวๆ (rule of thumb) เช่น 5% ของประชากร 10 ตัวอย่างต่อ 1 ตัวแปร

10. การเลือกตัวอย่าง หมายถึง การเลือกตัวอย่างแบบใช้หลักความน่าจะเป็น (probability sampling) หรือไม่ใช้หลักความน่าจะเป็น (nonprobability sampling) ดังนี้

10.1 การเลือกตัวอย่างแบบใช้หลักความน่าจะเป็น ได้แก่ การเลือกตัวอย่างแบบสุ่มอย่างง่าย (simple random sampling) การเลือกตัวอย่างแบบสุ่มอย่างเป็นระบบ (systematic random sampling) การเลือกตัวอย่างแบบสุ่มด้วยการแบ่งชั้น/ชั้นภูมิ (stratified random sampling) และการเลือกตัวอย่างแบบสุ่มเชิงพื้นที่ (cluster random sampling)

10.2 การเลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น ได้แก่ การเลือกตัวอย่างแบบสัดส่วนหรือโควต้า (quota sampling) การเลือกตัวอย่างแบบตามสะดวก/เผอิญพบ (convenience/accidental sampling) การเลือกตัวอย่างแบบเจาะจง (judgmental/purposive sampling)

11. การวิเคราะห์ข้อมูล หมายถึง การใช้สถิติแบบมีพารามิเตอร์ (parametric statistics) หรือไม่มีพารามิเตอร์ในการวิเคราะห์ข้อมูล (nonparametric statistics) ดังนี้

11.1 สถิติแบบมีพารามิเตอร์ ได้แก่ การทดสอบค่าเฉลี่ยกลุ่มตัวอย่าง 1 กลุ่มด้วยค่าที (One-sample t-test) การทดสอบสอบค่าเฉลี่ยระหว่าง 2 กลุ่มตัวอย่างด้วยค่าที (Two-sample t-test) การทดสอบค่าเฉลี่ยระหว่างกลุ่มตัวอย่างที่มีความสัมพันธ์กันด้วยค่าที (Paired t-test) การวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA) การวิเคราะห์ความแปรปรวนแบบสองทาง (Two-way ANOVA) การวิเคราะห์ความแปรปรวนแบบวัดซ้ำ (Repeated Measures ANOVA) การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) การวิเคราะห์การถดถอยเชิงเส้นตรง (Linear Regression) การวิเคราะห์องค์ประกอบ (Factor Analysis) และการวิเคราะห์ตัวแปรพหุ (Multivariate Analysis)

11.2 สถิติแบบไม่มีพารามิเตอร์ ได้แก่ การทดสอบความสอดคล้องของข้อมูลด้วยไคสแควร์ (X^2 goodness of fit) การทดสอบความเป็นอิสระของข้อมูลด้วยไคสแควร์ (X^2 Test for Independence) การทดสอบค่ากลางของข้อมูล 1 กลุ่มตัวอย่าง (One-sample Test) การทดสอบค่ากลางของกลุ่มตัวอย่างที่มีความสัมพันธ์กัน (2 Related Samples Test) การทดสอบค่ากลางของกลุ่มตัวอย่างที่เป็นอิสระจากกัน (Independence Sample Test) การทดสอบสหสัมพันธ์แบบสเปียร์แมน (Spearman Correlation Test) การวิเคราะห์การถดถอยแบบไม่มีพารามิเตอร์ (Nonparametric Regression Analysis) การวิเคราะห์พหุแบบไม่มีพารามิเตอร์ (Nonparametric Multivariate Analysis) และรวมถึงสถิติเชิงพรรณนา

12. สมมติฐาน หมายถึง มีการกำหนดสมมติฐานหรือไม่มีการกำหนดสมมติฐานเพื่อใช้ในการทดสอบด้วยวิธีการทางสถิติ

บทที่ 3
ระเบียบวิธีการวิจัย

การวิจัยนี้ใช้ระเบียบวิธีการวิจัยเชิงปริมาณ (quantitative methodology) รวบรวมข้อมูลจากวิทยานิพนธ์และดุษฎีนิพนธ์ (thesis and dissertation) สาขาสังคมวิทยา จากฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทย (ThaiLIS-Thai Library Integrated System) ของสำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ[1] โดยมีรายละเอียดดังนี้

การพัฒนากรอบแนวคิดการวิจัย

ทบทวนวรรณกรรมเกี่ยวกับหลักการออกแบบการวิจัย การวิจัยเชิงปริมาณ ความคลาดเคลื่อนในการวิจัย และการวิเคราะห์ความเสี่ยง เพื่อนำมาใช้ในการสร้างตัวแบบในการวิเคราะห์และการวัดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย

หน่วยวิเคราะห์

หน่วยที่ใช้ในการวิเคราะห์ (unit of analysis) เป็นวิทยานิพนธ์และดุษฎีนิพนธ์ สาขาสังคมวิทยา ที่ใช้วิธีการวิจัยเชิงปริมาณ

เครื่องมือการวิจัย

เครื่องมือที่ใช้ในการวิจัยเป็นแฟ้มข้อมูล (file) แผ่นตารางทำการ (spread sheet) สร้างเป็นแบบบันทึกข้อมูลเพื่อรองรับการรวบรวมข้อมูลจากสืบค้นข้อมูลด้วยโปรแกรมค้นหาข้อมูล (search engine) การแปลงข้อมูลด้วยภาษา XML (eXtensible Markup Language) และการบันทึกข้อมูลจากแบบฟอร์มที่สร้างด้ายโปรแกรมคอมพิวเตอร์ เช่น AppSheet, Google Form โดยมีรายละเอียด ดังนี้

[1] ปัจจุบันสังกัดกระทรวงการอุดมศึกษา วิทยาศาสตร์ วิจัยและนวัตกรรม

ส่วนที่ 1 ข้อมูลทั่วไปเกี่ยวกับวิทยานิพนธ์และดุษฎีนิพนธ์ เป็นข้อมูลที่รวบรวมครั้งที่ 1 เพื่อนำมาใช้คัดกรองประเภทงานวิจัยและกำหนดกรอบการสุ่มตัวอย่าง มีตัวแปรที่เกี่ยวข้องดังนี้

1. ลำดับที่ (No.) คือ ตัวเลขที่กำหนดให้เป็นลำดับที่ของวิทยานิพนธ์หรือดุษฎีนิพนธ์ที่รวบรวมมาจากฐานข้อมูลฯ

2. เลขเรียกหนังสือ (Call number) คือ เลขรหัสของวิทยานิพนธ์และดุษฎีนิพนธ์แต่ละเรื่อง/เล่มที่ห้องสมุดของแต่ละมหาวิทยาลัยกำหนดไว้สำหรับใช้ค้นหาตัวเล่มจากห้องสมุด

3. ปี (Year) คือ ปี พ.ศ. ที่วิทยานิพนธ์และดุษฎีนิพนธ์แล้วเสร็จ

4. ชื่อปริญญา (Degree name) คือ ชื่อปริญญาของแต่ละหลักสูตร

5. ระดับปริญญา (Level) คือ ระดับปริญญาของแต่ละหลักสูตร

6. สาขาวิชา (Discipline) คือ ชื่อกลุ่มสาขาวิชาตามการจัดหมวดหมู่ของฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทยของสำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ

7. มหาวิทยาลัย (University) คือ มหาวิทยาลัยที่เป็นเจ้าของวิทยานิพนธ์และดุษฎีนิพนธ์

8. ประเภทงานวิจัย (Research type) คือ เป็นวิทยานิพนธ์และดุษฎีนิพนธ์ที่ใช้การวิจัยเชิงปริมาณ เชิงคุณภาพ หรือเชิงผสมผสาน

ส่วนที่ 2 ข้อมูลเกี่ยวกับระเบียบวิธีการวิจัย เป็นข้อมูลที่รวบรวมเพิ่มเติมรอบที่ 2 หลังจากการสุ่มตัวอย่างแล้ว เพื่อนำไปใช้รวบรวมข้อมูลจากวิทยานิพนธ์หรือดุษฎีนิพนธ์แต่ละเรื่อง โดยใช้ Google Form สร้างเป็นแบบสำรวจ มีข้อคำถามดังต่อไปนี้

1. เลขที่ของวิทยานิพนธ์หรือดุษฎีนิพนธ์ (No.) เป็นข้อคำถามแบบปลายเปิดประเภทคำตอบข้อความแบบสั้น (short answer text) เพื่อนำเอาเลขลำดับที่ของวิทยานิพนธ์หรือดุษฎีนิพนธ์ที่กำหนดไว้ในส่วนที่ 1 มาใส่ สำหรับใช้ในการเชื่อมโยงข้อมูลและค้นหาข้อมูลในส่วนที่ 1 ดังนี้

2. มีการกำหนดสมมติฐานเพื่อการทดสอบสมมติฐาน (Hypothesis) เป็นคำถามแบบปลายปิดประเภทเลือกตอบ (multiple choice) ดังนี้

มีการกำหนดสมมติฐานเพื่อการทดสอบสมมติฐาน (hypothesis) *

- ○ มี
- ○ ไม่มี

3. การสร้างเครื่องมือวัด (Measurement design) เป็นคำถามแบบปลายปิดประเภทตารางมีช่องให้เลือกตอบได้หลายข้อ (checkbox grid) ดังนี้

การสร้างเครื่องมือวัด (measurement design)

	มี
มีการทดสอบความเที่ยงตรง (validity)	☐
มีการทดสอบความเชื่อมั่น (reliability)	☐

4. ประเภทประชากร (Population type) เป็นคำถามแบบปลายปิดประเภทเลือกตอบ (multiple choice) ดังนี้

ประเภทประชากร (population type) *

- ○ ทราบจำนวนประกร (finite population)
- ○ ไม่ทราบจำนวนประชากร (infinite population)

5 ขนาดตัวอย่าง (Sample size) เป็นคำถามแบบปลายปิดประเภทเลือกตอบ (multiple choice) ดังนี้

ขนาดตัวอย่าง (sample size) *

- ○ คำนวณด้วยวิธีการกำหนดความช่วงความเชื่อมั่น (confidence interval)
- ○ คำนวณด้วยวิธีการกำหนดขนาดผลกระทบ (effect size)
- ○ Other...

6. ประเภทการเลือกตัวอย่าง (Sampling type) เป็นคำถามแบบปลายปิดประเภทเลือกตอบ (multiple choice) ดังนี้

ประเภทการเลือกตัวอย่าง (sampling type) *

- ○ เลือกตัวอย่างแบบใช้หลักความน่าจะเป็น (probability sampling)
- ○ เลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น (nonprobability sampling)
- ○ เลือกตัวอย่างแบบผสม (combination sampling)

7. วิธีเลือกตัวอย่าง (Sampling method) เป็นคำถามแบบปลายปิดประเภทตารางมีช่องให้เลือกตอบได้หลายข้อ (Checkbox grid) ดังนี้

วิธีเลือกตัวอย่าง (sampling method)

	ใช่
การเลือกตัวอย่างแบบสุ่มอย่างง่าย (simple random sampling)	☐
การเลือกตัวอย่างแบบสุ่มด้วยการแบ่งชั้น/ชั้นภูมิ (stratified rand...	☐
การเลือกตัวอย่างแบบสุ่มอย่างเป็นระบบ (systematic random s...	☐
การเลือกตัวอย่างแบบสุ่มเชิงพื้นที่ (cluster random sampling)	☐
การเลือกตัวอย่างแบบสัดส่วนหรือโควต้า (quota sampling)	☐
การเลือกตัวอย่างแบบเผอิญพบ (convenience/accidental sam...	☐
การเลือกตัวอย่างแบบเจาะจง (judgmental/purposive sampling)	☐
การเลือกตัวอย่างเชิงก้อนหิมะหรือลูกโซ่ (snowball/chain referr...	☐

8. ระดับตัวแปร (Variable level) เป็นคำถามแบบปลายปิดประเภทตารางมีช่องให้เลือกตอบได้หลายข้อ (checkbox grid) ดังนี้

ระดับตัวแปรวัด (variable level)

	ใช่
นามมาตรา (nominal scale)	☐
อันดับมาตรา (ordinal scale)	☐
ช่วงมาตรา (interval scale)	☐
สัดส่วนมาตรา (ratio scale)	☐

9. การตรวจสอบค่าสูญหาย (Missing detected) เป็นคำถามแบบปลายปิดประเภทเลือกตอบ (multiple choice) ดังนี้

การตรวจสอบค่าสูญหาย (detecting missing values) *

- ○ ตรวจสอบ
- ○ ไม่ตรวจสอบ

10. การวิเคราะห์ข้อมูลเบื้องต้น (Basic analysis) เป็นคำถามแบบปลายปิดประเภทตารางมีช่องให้เลือกตอบได้หลายข้อ (checkbox grid) ดังนี้

การวิเคราะห์ข้อมูลเบื้องต้น (basic analysis)

	ใช่
การตรวจสอบการกระจายหรือแจกแจงแบบปกติของข้อมูล (norm...	☐
การตรวจสอบความเท่ากันของความแปรปรวนระหว่างกลุ่ม (homo...	☐
การตรวจสอบความเท่ากันของความแปรปรวนของความคลาดเคลื่...	☐
การตรวจสอบความเท่ากันของความแปรปรวน-ความแปรปรวนร่วม...	☐
การตรวจสอบความเป็นเส้นตรง (linearity)	☐
การตรวจสอบภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรอิสระหลายตัวแป...	☐
การตรวจสอบความสัมพันธ์หรือความเป็นอิสระภายในตัวแปร (aut...	☐
การตรวจสอบความสุ่ม (randomness)	☐

11. การปรับแก้ข้อมูล (Data transformations) เป็นคำถามแบบปลายปิดประเภทเลือกตอบ (multiple choice) ดังนี้

การปรับแก้ข้อมูล (data transformations) *

- ○ มีการปรับแก้ข้อมูล
- ○ ไม่มีการปรับแก้ข้อมูล

12. ประเภทสถิติ (Statistic type) เป็นคำถามแบบปลายปิดประเภทตารางมีช่องให้เลือกตอบได้หลายข้อ (checkbox grid) แบ่งเป็น 3 ข้อ ดังนี้

สถิติเชิงอ้างอิง (inference statistics) แบบมีพารามิเตอร์ (parametric tests)

	ใช่
การทดสอบค่าเฉลี่ยของกลุ่มตัวอย่าง 1 กลุ่ม (one-sample...	☐
การทดสอบค่าเฉลี่ยแบบ 2 กลุ่ม (two-sample t-test)	☐
การทดสอบค่าเฉลี่ยแบบ 2 กลุ่ม ไม่เป็นอิสระจากกัน (paire...	☐
การทดสอบค่าเฉลี่ยแบบ 3 กลุ่ม (one-way ANOVA)	☐
การทดสอบความแปรปรวนแบบ 2 ทาง (two-way ANOVA)	☐
การวิเคราะห์ความแปรปรวนที่มีการวัดซ้ำ (repeated meas...	☐
การทดสอบด้วยสัมประสิทธิ์ความสัมพันธ์เพียร์สัน (pearso...	☐
การทดสอบด้วยการวิเคราะห์ถดถอยเชิงเส้นตรงแบบง่าย (s...	☐
การทดสอบด้วยการวิเคราะห์ถดถอยเชิงเส้นตรงแบบพหุคุณ...	☐

สถิติเชิงอ้างอิง (inference statistics) แบบไม่มีพารามิเตอร์ (non-parametric tests)

	ใช่
การทดสอบแบบไคสแควร์ (chi-square test)	☐
การทดสอบด้วยสัมประสิทธิ์สหสัมพันธ์สเปียร์แมนแรงค์ (sp...	☐
การทดสอบด้วยวิธีของแมน-วิทนี ยู (Mann-Whitney U test)	☐
การทดสอบด้วยวิธีของครัสคัล-วัลลิส (Kruskal-Wallis H te...	☐
การทดสอบด้วยวิธีของวิลคอกซัน (Wilcoxon test	☐
การทดสอบด้วยวิธีของฟลิดแมน (Friedman test)	☐
การทดสอบด้วยการวิเคราะห์การถดถอยโลจิสติก (logistic ...	☐
การทดสอบด้วยการวิเคราะห์การถดถอยพหุโลจิสติกส (mul...	☐

สถิติแบบอื่นๆ

	ใช่
การวิเคราะห์เส้นทาง (path analysis)	☐
การวิเคราะห์สมการโครงสร้าง (structural equation mod...	☐
การวิเคราะห์องค์ประกอบ (factor analysis)	☐
การวิเคราะห์จำแนกประเภท (discriminant analysis)	☐
การวิเคราะห์การจัดกลุ่ม (cluster analysis)	☐
การวิเคราะห์อนุกรมเวลา (time series analysis)	☐
การวิเคราะห์การอยู่รอด (survival analysis)	☐
การวิเคราะห์หมวดหมู่ (classification)	☐

13. หมายเหตุ (Note) เป็นข้อคำถามแบบปลายเปิดประเภทคำตอบข้อความแบบยาวหรือพรรณนา (long answer text) เพื่อบันทึกข้อมูลเพิ่มเติมเกี่ยวกับวิทยานิพนธ์หรือดุษฎีนิพนธ์แต่ละเรื่อง ดังนี้

หมายเหตุ

Long answer text

ประชากร

ประชากรเป้าหมาย (target population) คือ วิทยานิพนธ์และดุษฎีนิพนธ์ สาขาสังคมวิทยา ที่ใช้วิธีการวิจัยเชิงปริมาณเพียงวิธีการเดียวเท่านั้น (ไม่รวมการวิจัยแบบผสมที่ใช้การวิจัยเชิงปริมาณและการวิจัยเชิงคุณภาพ) จากหลักสูตรปริญญาโทหรือปริญญาเอก สาขาสังคมวิทยา ที่อยู่ในฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทยของสำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ ที่แล้วเสร็จระหว่างปี พ.ศ. 2512-2560

การรวบรวมข้อมูลรายชื่อวิทยานิพนธ์และดุษฎีนิพนธ์

1. การสำรวจข้อมูลเบื้องต้น เมื่อวันที่ 5 สิงหาคม พ.ศ. 2561 ใช้เครื่องมือค้นข้อมูล (search engine) แบบ Advanced Search จากเว็บไซต์ http://tdc.thailis.or.th/tdc/ สืบค้นข้อมูลด้วยคำว่า สังคมวิทยา โดยกำหนดเงื่อนไขเพิ่มในการค้นข้อมูล ดังนี้

วิธีค้นข้อมูล = ส่วนใดส่วนหนึ่ง

มหาวิทยาลัย/สถาบัน = ทุกมหาวิทยาลัย/สถาบัน

เขตข้อมูล = สาขาวิชา

ชนิดเอกสาร = วิทยานิพนธ์/Thesis

ส่วนเงื่อนไขอื่น ไม่กำหนด คือ จำกัดข้อมูลเฉพาะ ผลการสืบค้นมีวิทยานิพนธ์สาขาสังคมวิทยาอยู่ในฐานข้อมูลในวันดังกล่าวจำนวน 1,171 เรื่อง

2. การส่งออกข้อมูลจากระบบฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทยในรูปของข้อมูลแบบ XML (eXtensible Markup Language) เป็นแฟ้มข้อมูล (files) แบบ xml แฟ้มข้อมูลละ 20 เรื่อง จำนวน 59 แฟ้มข้อมูล

2.1 การสร้างโครงร่างรับข้อมูล (XML schema) จากแฟ้มข้อมูลส่งออกสำหรับนำเข้าโปรแกรมตารางคำนวณ ดังนี้

<?xml version="1.0" encoding="TIS-620"?>

<SociologyThesis>

<TDC>

<Title></Title>
<TitleAlternative></TitleAlternative>
<Creator></Creator>
<DateCreated></DateCreated>
<Type></Type>
<SourceCallNumber></SourceCallNumber>
<ThesisDegreeName></ThesisDegreeName>
<ThesisLevel></ThesisLevel>
<ThesisDescipline></ThesisDescipline>
<ThesisGrantor></ThesisGrantor>
</TDC>
<TDC>
<Title></Title>
<TitleAlternative></TitleAlternative>
<Creator></Creator>
<DateCreated></DateCreated>
<Type></Type>
<SourceCallNumber></SourceCallNumber>
<ThesisDegreeName></ThesisDegreeName>
<ThesisLevel></ThesisLevel>
<ThesisDescipline></ThesisDescipline>
<ThesisGrantor></ThesisGrantor>
</TDC>
</SociologyThesis>

2.2 การแปลงข้อมูล XML เข้าโปรแกรมตารางคำนวณ (spreadsheet) ของ Microsoft Excel โดยสร้างเป็นชุดข้อมูล (data set) ในรูปของตาราง (table) ประกอบด้วยสดมภ์ (column) ของตัวแปร (variable) และแถว (row) ของชุดข้อมูล (case)

2.3 ตรวจสอบข้อมูล (data cleansing) ปรับแก้ข้อมูลที่สะกดผิดหรือใช้ชื่อแตกต่างไปจากคำที่เป็นชื่อหลักๆ เช่น ชื่อปริญญา สาขาวิชา และชื่อมหาวิทยาลัย รวมถึงการลงปี พ.ศ. ที่แล้วเสร็จไม่ตรงกับ

หน้าปกวิทยานิพนธ์ เช่น ลงเป็นปี ค.ศ. ระหว่างการตรวจสอบข้อมูลพบว่า มีข้อมูลของวิทยานิพนธ์ฉบับหนึ่งเป็นวิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต แต่ลงข้อมูลเป็นสาขาวิชาสังคมวิทยา จึงตัดวิทยานิพนธ์ฉบับดังกล่าวออกจากชุดข้อมูล (dataset) ทำให้เหลือจำนวนวิทยานิพนธ์ในการจำแนกประเภทวิทยานิพนธ์ทั้งหมด 1,170 เรื่อง

การจำแนกประเภทวิทยานิพนธ์และดุษฎีนิพนธ์

เนื่องจากต้องระบุประเภทวิทยานิพนธ์และดุษฎีนิพนธ์จำนวน 1,170 เรื่อง ต้องใช้ผู้ช่วยนักวิจัยหลายคน จึงใช้นิสิตระดับปริญญาตรีที่เรียนวิชา ระเบียบวิธีการวิจัยทางสังคมศาสตร์ ที่ผู้วิจัยเป็นผู้สอน ทำการจำแนกประเภทงานวิจัยเรื่องใดเป็นการวิจัยเชิงปริมาณหรือไม่ใช่การวิจัยเชิงปริมาณ โดยมีกระบวนการดังนี้

1. การแบ่งปันข้อมูล นำชุดข้อมูลจากแฟ้มข้อมูลโปรแกรมของ Microsoft Excel เข้าโปรแกรมตารางคำนวณ Google Sheet เพื่อให้แบ่งปัน (share) ข้อมูลให้ผู้ช่วยนักวิจัยทำการสำรวจข้อมูลประเภทของงานวิจัยและตรวจสอบปีที่จัดทำเอกสารตัวเล่มจากงานวิจัยแต่ละเล่ม

2. การทำแอป (apps) ด้วยโปรแกรมเสริมขนาดเล็ก (add-on) ชื่อ AppSheet ในระบบ Google Sheet เพื่อดึงข้อมูลวิทยานิพนธ์แต่ละเล่มมาใส่ข้อมูลของตัวแปรประเภทวิทยานิพนธ์ รวมถึงตรวจสอบและแก้ไขปีที่จัดทำเอกสารตัวเล่มให้ถูกต้อง

3. การแนะนำผู้ช่วยนักวิจัยในการจำแนกประเภทวิทยานิพนธ์ การใช้แอป และการค้นข้อมูลจากฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทย

4. การค้นหาวิทยานิพนธ์ ใช้ชื่อผู้ทำวิทยานิพนธ์ค้นข้อมูลจากฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทยทีละเล่ม หากได้ชื่อวิทยานิพนธ์ตรงกับชื่อผู้ทำวิทยานิพนธ์ จึงเริ่มการจำแนกประเภทวิทยานิพนธ์

5. การจำแนกประเภทวิทยานิพนธ์ พิจารณาจากดัชนีที่บ่งบอกถึงประเภทการวิจัย เช่น วิธีการวิจัย การเก็บข้อมูล จำนวนตัวอย่าง และวิธีการวิเคราะห์ข้อมูล โดยมีวิธีการค้นหาข้อมูล ดังนี้

ขั้นที่ 1 เริ่มจากหน้าแสดงรายละเอียดเกี่ยวกับวิทยานิพนธ์ (metadata) โดยดูจากบทคัดย่อ (abstract) ในหัวข้อ Description

ขั้นที่ 2 กรณีที่ Description ในหน้าแสดงรายละเอียดไม่มีข้อมูลบทคัดย่อ ให้เปิดบทคัดย่อที่เป็นแฟ้มข้อมูลแบบ pdf จากรายการแฟ้มข้อมูล

ขั้นที่ 3 กรณีในบทคัดย่อมีข้อมูลไม่เพียงพอในการตัดสินใจ ให้เปิดแฟ้มข้อมูลแบบ pdf บทที่ 3 หรือบทที่ 2 หรือบทที่ 1 ตามลำดับ เพราะวิทยานิพนธ์แต่ละช่วงเวลาเขียนวิธีการวิจัยไว้ในแต่ละบทแตกต่างกัน

ขั้นที่ 4 กรณีไม่สามารถเปิดแฟ้มข้อมูลแบบ pdf จากฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทยได้ เข้าไปค้นหาจากเว็บไซต์ของห้องสมุดของแต่ละมหาวิทยาลัย

จากการจำแนกประเภทวิทยานิพนธ์ พบวิทยานิพนธ์ที่ใช้วิธีการวิจัยเชิงปริมาณ จำนวน 573 เรื่อง

กลุ่มตัวอย่าง

การสร้างกรอบการสุ่มตัวอย่าง (sampling frame) เริ่มจากการรวบรวมรายละเอียดหรือคำอธิบาย (metadata) เกี่ยวกับวิทยานิพนธ์และดุษฎีนิพนธ์ ได้แก่ ชื่อเรื่อง ชื่อผู้จัดทำ ปีที่แล้วเสร็จ ชื่อปริญญา ระดับปริญญา สาขาวิชา และมหาวิทยาลัย จากฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทยของสำนักงานคณะกรรมการการอุดมศึกษา ที่เผยแพร่ผ่านเว็บไซต์ http://tdc.thailis.or.th/tdc/ โดยมีขั้นดังนี้

การกำหนดขนาดตัวอย่าง

ตัวแปรวัดในการวิจัยใช้มาตรวัดแบบสองคำตอบ (dichotomy/binary) มีค่าเป็น 0 = ไม่ใช่ และ 1 = ใช่ จึงใช้วิธีการกำหนดขนาดกลุ่มตัวอย่าง (sample size) ด้วยสูตรของสูตรทาโร่ ยามาเน่ (Yamane, 1967: 100) ดังนี้

$$n = \frac{N}{Nd^2 + 1}$$

เมื่อ n = ขนาดกลุ่มตัวอย่าง

N = ขนาดประชากร

d^2 = ความคลาดเคลื่อนของกลุ่มตัวอย่าง

โดยกำหนดค่าช่วงความเชื่อมั่น (confidence interval) ที่ 95% หรือระดับความแม่นยำ (precision) ที่ 5%

แทนค่า $\frac{573}{(573X0.0025)+1}$ = 235.56

ขนาดตัวอย่างที่ได้จากการคำนวณเป็นวิทยานิพนธ์จำนวน 236 เรื่อง และเพื่อป้องกันการไม่สามารถเข้าถึงแฟ้มข้อมูลและเปิดแฟ้มข้อมูลวิทยานิพนธ์ที่เป็นตัวอย่างได้ เช่น แฟ้มข้อมูลเสีย ไม่มีแฟ้มข้อมูล จึงเตรียมจำนวนตัวอย่างเพิ่มไว้ 10% หรือ 22 เรื่อง รวมจำนวนตัวอย่างทั้งหมดที่ใช้ในการรวบรวมข้อมูล 243 เรื่อง

การเลือกตัวอย่าง

เลือกตัวอย่างแบบอาศัยความน่าจะเป็น (probability sampling) โดยใช้วิธีการสุ่มแบบง่าย (simple random sampling) ดังนี้

1. นำรายชื่อวิทยานิพนธ์จำนวน 573 เรื่อง มาใส่เลขสุ่ม (random number) ด้วยคำสั่ง rand()

2. เรียงลำดับ (order) เลขสุ่ม ด้วยคำสั่งเริ่มต้น (default) ที่เรียงจากค่าน้อยไปหาค่ามาก (smallest to largest)

3. คัดเลือกวิทยานิพนธ์และดุษฎีนิพนธ์จำนวน 236 เรื่อง จากชุดข้อมูลรายชื่อวิทยานิพนธ์ 573 เรื่อง เพื่อใช้ในการรวบรวมข้อมูลตามรายการตัวแปร

การรวบรวมข้อมูล

นำรายชื่อวิทยานิพนธ์และดุษฎีนิพนธ์ที่ได้จากกระบวนการเลือกตัวอย่างไปสืบค้นข้อมูลจากฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทย โดยใช้ผู้ช่วยนักวิจัยที่จบการศึกษาระดับปริญญาโท เปิดแฟ้มข้อมูลตรวจหาข้อมูลตามรายการตัวแปรวัดจากเนื้อหาในบทต่างๆ และบันทึกข้อมูลตามแบบฟอร์มสำรวจข้อมูลที่สร้างด้วย Google form

การวิเคราะห์ข้อมูล

ใช้วิธีการวิเคราะห์หลากหลายวิธี ทั้งสถิติเชิงพรรณา คือ ค่าความถี่ และค่าร้อยละ และสถิติเชิงอ้าอิง คือ การวิเคราะห์จัดกลุ่ม และการวิเคราะห์การถดถอยแบบโลจิสติก โดยใช้โปรแกรมสำเร็จรูปทางสถิติ (SPSS และ R) มีรายละเอียดดังต่อไปนี้

1. การวิเคราะห์ข้อมูลทั่วไปเกี่ยวกับวิทยานิพนธ์และดุษฎีนิพนธ์ ใช้สถิติพรรณนา (descriptive statistics) ได้แก่ การแจกแจงความถี่ (frequency) และค่าร้อยละ (percentage)

2. การวัดและการให้คะแนนดัชนีความเสี่ยง นำข้อมูลที่ได้จากการวบรวมในรอบที่ 2 มาสร้างตัวแปรหุ่น (dummy variables) และกำหนดค่าคะแนนใหม่เป็นแบบสองค่า (binary data) เพื่อนำไปใช้ในการวิเคราะห์หาความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย โดยมีค่า (value) และความหมาย (label) ดังรายละเอียดตามตารางที่ 3.1

ตารางที่ 3.1 แสดงค่าและความหมายของตัวแปรหุ่น

ตัวแปรหุ่น	ค่าตัวแปรและความหมาย	ค่าตัวแปรจากตัวแปรวัด
สมมติฐาน	0 = ไม่มีสมติฐาน	ไม่มีสมมติฐาน
	1 = มีสมมุติฐาน	มีสมมติฐาน
สถิติ	0 = สถิติแบบไม่มีพารามิเตอร์	สถิติเชิงพรรณนา และสถิติเชิงอ้างอิงแบบไม่มีพารามิเตอร์
	1 = สถิติแบบมีพารามิเตอร์	สถิติเชิงอ้างอิงแบบมีพารามิเตอร์
เครื่องมือวัด	0 = ไม่มีการทดสอบ	ไม่มีการทดสอบความเที่ยงตรงและไม่มีการทดสอบสอบความเชื่อมั่น
	1 = มีการทดสอบ	มีการทดสอบความเที่ยงตรงและ/หรือมีการทดสอบความเชื่อมั่น
ระดับการวัด	0 = ค่าไม่ต่อเนื่อง	การวัดแบบนามมาตรา และการวัดแบบอันดับมาตรา
	1 = ค่าต่อเนื่อง	การวัดแบบช่วงมาตรา และการวัดแบบสัดส่วนมาตรา

ตัวแปรหุ่น	ค่าตัวแปรและความหมาย	ค่าตัวแปรจากตัวแปรวัด
การตรวจสอบข้อมูล	0 = ไม่มีการตรวจสอบ	ไม่ตรวจสอบค่าสูญหาย ไม่ตรวจการแจกแจงปกติ และไม่ตรวจสอบข้อมูลตามข้อตกลงเบื้องเบื้องต้น
	1 = มีการตรวจสอบ	ตรวจสอบค่าสูญหาย หรือตรวจการแจกแจงปกติ และ/หรือตรวจสอบข้อมูลตามข้อตกลงเบื้องต้น
ประชากร	0 = ไม่ทราบจำนวน	ไม่ทราบจำนวนประชากร
	1 = ทราบจำนวน	ทราบจำนวนประชากร
ขนาดตัวอย่าง	0 = ไม่คำนวนด้วยวิธีการทางสถิติ	สำมะโน เจาะจง งานวิจัยไม่ได้กล่าวถึง
	1 = คำนวนด้วยวิธีการทางสถิติ	คำนวณด้วยวิธีการกำหนดความช่วงความเชื่อมั่น
การเลือกตัวอย่าง	0 = ไม่ใช้หลักความน่าจะเป็น	การเลือกตัวอย่างแบบสัดส่วนหรือโควตา การเลือกตัวอย่างแบบตามสะดวก/เผอิญพบ การเลือกตัวอย่างแบบเจาะจง
	1 = ใช้หลักความน่าจะเป็น	การเลือกตัวอย่างแบบสุ่มอย่างง่าย การเลือกตัวอย่างแบบสุ่มอย่างเป็นระบบ การเลือกตัวอย่างแบบสุ่มด้วยการแบ่งชั้น/ชั้นภูมิ และการเลือกตัวอย่างแบบสุ่มเชิงพื้นที่

3. การวิเคราะห์จัดกลุ่ม (Cluster Analysis: CA) เนื่องจากมีวิทยานิพนธ์กลุ่มตัวอย่างจำนวนมาก (มากกว่า 200 เรื่อง) และมีการดำเนินการตามกระบวนการวิจัยที่แตกต่างกันมาก ทำให้ไม่ทราบว่ามีการออกแบบการวิจัยตามระเบียบวิธีการวิจัยเชิงปริมาณตามองค์ประกอบของกรอบแนวคิดการวิจัยกี่รูปแบบ จึงใช้การวิธีการวิเคราะห์จัดกลุ่มแบบผสม (combination algorithm) ทั้งแบบเป็นขั้นตอน (hierarchical cluster analysis) และแบบไม่เป็นขั้นตอน (nonhierarchical cluster analysis) แบ่งกลุ่มวิทยานิพนธ์เป็นกลุ่มๆ เพื่อให้ได้รูปแบบการดำเนินการวิจัยที่เหมือนกัน โดยมีขั้นตอนดังนี้

3.1 การวิเคราะห์จัดกลุ่มแบบสองขั้นตอน (Two Step Cluster Analysis: TSCA) เพื่อหาจำนวนกลุ่มที่เหมาะสมเพื่อนำไปใช้ในการจัดกลุ่มแบบเคมีน (k-means clustering) โดยใช้การวัดระยะทางแบบ

ความควรจะเป็นเชิงลอการิทึม (log likelihood) และพิจารณาค่าคุณภาพการจัดกลุ่ม (cluster quality) จากค่าเฉลี่ยสัมประสิทธิ์ซิลลูเอท (silhouette coefficient) มีขั้นตอนวิธี (algorithm) ดังนี้

ขั้นที่ 1 ก่อนการจัดกลุ่ม (Pre-clustering) แบ่งข้อมูลเป็นกลุ่มย่อย (divisive method) หรือเรียกว่า วิธีการจากบนลงล่าง (top-down approach) โดยเริ่มจากแบ่งข้อมูลทั้งหมดที่มีลักษณะ (attribute) หรือรูปแบบ (pattern) ที่เหมือนกันหรือคล้ายกัน (similarity) รวมเข้าไว้เป็นกลุ่มเดียวกัน ส่วนข้อมูลที่มีลักษณะหรือรูปแบบที่แตกต่างกันจะถูกจัดรวมให้อยู่ต่างกลุ่มกัน

ขั้นที่ 2 การจัดกลุ่ม (Clustering) ใช้วิธีจัดกลุ่มแบบเป็นขั้นตอน (hierarchical methods) รวมกลุ่มเข้า (agglomerative method) หรือเรียกว่า วิธีการจากล่างขึ้นบน (bottom-up approach) โดยเริ่มจากรวมกลุ่มย่อยๆ ที่เหมือนกันและรวมกลุ่มย่อยๆ ที่แตกต่างกันเข้าเป็นกลุ่มใหญ่ตามจำนวนกลุ่มที่ต้องการ

3.2 การวิเคราะห์จัดกลุ่มแบบเคมีน (K-means Cluster Analysis) เพื่อหารูปแบบของวิธีการวิจัยที่มีเหมือนกันและแตกต่างกัน มีขั้นตอนวิธี (algorithm) ดังนี้

ขั้นที่ 1 แบ่งกลุ่มข้อมูลทั้งหมดออกเป็น k กลุ่ม (ตามจำนวนกลุ่มที่ได้จากการวิเคราะห์แบ่งกลุ่มแบบสองขั้นตอน)

ขั้นที่ 2 คำนวณระยะห่าง (Distance: d) จากจุดศูนย์กลางของแต่ละกลุ่ม โดยใช้การวัดระยะทางแบบยุคลิด (Euclidean Distance: $D_{Euclidean}$) ดังนี้

$$D_{Euclidean} = \sqrt{(x_1 - y_1)(x_1 - y_1)^2 + (x_2 - y_2)(x_2 - y_2)^2 + \cdots + (x_n - y_n)(x_n - y_n)^2}$$

4. การวิเคราะห์ความเสี่ยงของการเกิดความคลาดเคลื่อน นำแบบจำลองหรือรูปแบบการดำเนินการวิจัยที่ได้จากการวิเคราะห์จัดกลุ่มมาวิเคราะห์กำหนดว่าวิทยานิพนธ์หรือดุษฎีนิพนธ์เรื่องใดมีความเสี่ยงและไม่มีความเสี่ยง โดยให้คะแนน 0 = ละเมิดข้อตกลงเบื้องต้น และ 1 = ตรงกับข้อตกลงเบื้องต้น ตามเงื่อนไขดังรายละเอียดในตารางที่ 3.2

ตารางที่ 3.2 แสดงเงื่อนไขของตัวแปรหุ่นทีมีผลต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย

ตัวแปรหุ่น	การละเมิดข้อตกลงเบื้องต้น			
	สถิติแบบมีพารามิเตอร์		สถิติแบบไม่มีพารามิเตอร์	
เครื่องมือวัด	0	1	0	1
ระดับการวัด	0	1	0	0

ตัวแปรหุ่น	การละเมิดข้อตกลงเบื้องต้น			
	สถิติแบบมีพารามิเตอร์		สถิติแบบไม่มีพารามิเตอร์	
การตรวจสอบข้อมูล	0	1	0	0
ประชากร	0	1	0	0
ขนาดตัวอย่าง	0	1	0	0
การเลือกตัวอย่าง	0	1	0	0
ความเสี่ยง	1 = มี 0 = ไม่มี			

5. การวิเคราะห์การทำนายการเกิดความเสี่ยง ใช้การวิเคราะห์การถดถอยแบบพหุโลจิสติก (Multiple Logistic Regression: MLR) เพื่อทดสอบสมการ (equation) ของตัวแปรในรูปของแบบจำลอง (model) การพยากรณ์หรือทำนาย (predicted) การเกิดความเสี่ยง (risk) โดยมีขั้นตอนดังนี้

5.1 การวิเคราะห์และทดสอบความสัมพันธ์ระหว่างตัวแปรทำนายกับตัวแปรความเสี่ยง เพื่อวิเคราะห์ความสัมพันธ์และตรวจสอบภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรอิสระ (Multicollinearity)

5.2 การวิเคราะห์ความเหมาะสมของแบบจำลอง ด้วยค่าความควรจะเป็น (likelihood value) โดยมีวิธีการดังต่อไปนี้

5.2.1 วิธีการนำตัวแปรเข้าแบบจำลองพร้อมกันทั้งหมด (Enter Method) เพื่อตรวจสอบความสอดคล้องของแบบจำลองโดยรวมของตัวแปรทำนาย 6 ตัวแปร คือ เครื่องมือวัด ระดับการวัด การตรวจสอบข้อมูล ประชากร ขนาดตัวอย่าง และการเลือกตัวอย่าง

5.2.2 วิธีการคัดเลือกตัวแปรทำนายออกจากแบบจำลองอย่างเป็นขั้นตอนโดยใช้ค่าสัดส่วนความควรจะเป็น (Backward Stepwise: Likelihood Ratio) เพื่อหาแบบจำลองที่เหมาะสมจากตัวแปรทำนาย 6 ตัวแปร คือ เครื่องมือวัด ระดับการวัด การตรวจสอบข้อมูล ประชากร ขนาดตัวอย่าง และการเลือกตัวอย่าง

5.2.3 วิธีการคัดเลือกตัวแปรทำนายเพิ่มเข้าแบบจำลองอย่างเป็นขั้นตอนโดยใช้ค่าสัดส่วนความควรจะเป็น (Forward Stepwise: Likelihood Ratio) เพื่อหาแบบจำลองที่มีตัวแปรทำนายที่ดีที่สุดจากตัวแปรทำนาย 6 ตัวแปร คือ เครื่องมือวัด ระดับการวัด การตรวจสอบข้อมูล ประชากร ขนาดตัวอย่าง และการเลือกตัวอย่าง

5.3 การวิเคราะห์การพยากรณ์หรือทำนาย (predicted) ความน่าจะเป็น (probability) ของโอกาสที่จะเกิด (odds) ความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย (risk) จากอัตราส่วนระหว่างมีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยกับไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัย โดยมีสมการในรูปของ odds หรือ odd ratio ดังนี้

$$odds = \frac{P(R)}{Q(R)} \text{ หรือ } e^{B_0+B_{x1}\ldots+B_{xn}}$$

$P(R)$ = โอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย

$Q(R)$ = โอกาสที่จะไม่มีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย

เนื่องจากการทำนายความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยด้วยตัวแปรทำนาย ต้องใช้สมการถดถอยเชิงเส้นตรง (linear regression model) จึงทำการแปลงสมการ odds ให้อยู่ในรูปของสมการ log ของ odds หรือ logit (p) ดังนี้

$$\log(odds) = B_0 + B_1X_1 \ldots + B_nX_n$$

ดังนั้นสมการในการทำนายความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยจึงเป็นดังนี้

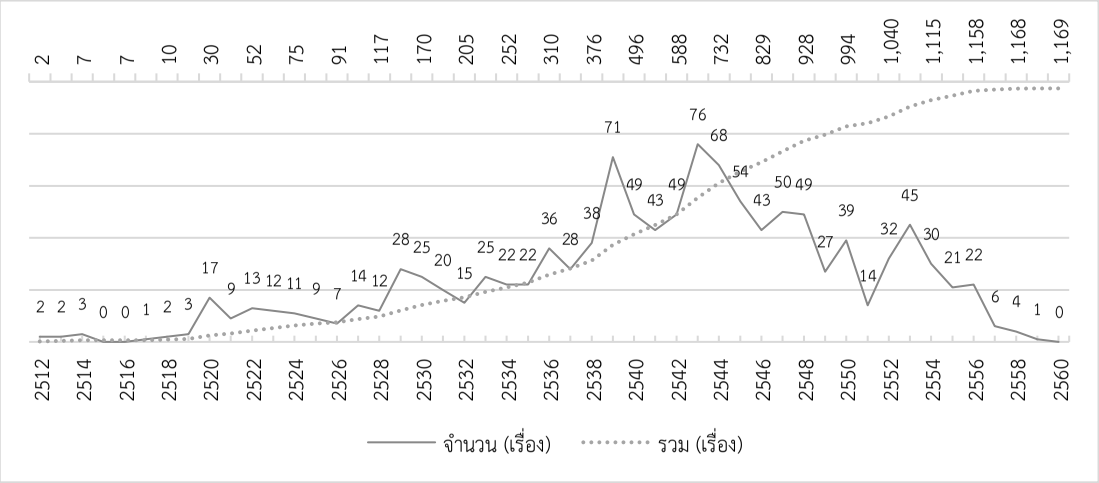
$$P(R) = \frac{e^{B_0+B_{x1}\ldots+B_{xn}}}{1+e^{B_0+B_{x1}\ldots+B_{xn}}} \text{ หรือ } \frac{1}{1+e^{-(B_0+B_1X_1\ldots+B_nX_n)}}$$

บทที่ 4
วิทยานิพนธ์และดุษฎีนิพนธ์สาขาวิชาสังคมวิทยา

การสำรวจวิทยานิพนธ์และดุษฎีนิพนธ์สาขาสังคมวิทยาจากฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศไทย สำนักงานคณะกรรมการการอุดมศึกษา เมื่อวันที่ 5 สิงหาคม พ.ศ. 2561 เพื่อใช้เป็นประชากรในการวิจัย (target population) ผู้วิจัยได้นำมาวิเคราะห์ให้เห็นภาพรวมของประชากรและแสดงให้เห็นพัฒนาการของงานวิจัยที่ผู้เรียนต้องจัดทำเพื่อเป็นส่วนหนึ่งของการศึกษาในระดับบัณฑิตศึกษาในสาขาวิชาสังคมวิทยา โดยแบ่งเนื้อหาออกเป็น 2 ส่วน คือ ข้อมูลทั่วไปเกี่ยวกับวิทยานิพนธ์และดุษฎีนิพนธ์ และข้อมูลทั่วไปเกี่ยวกับวิทยานิพนธ์และดุษฎีนิพนธ์ที่เป็นงานวิจัยเชิงปริมาณ (sampling frame) โดยมีรายละเอียดดังต่อไปนี้

ข้อมูลทั่วไปเกี่ยวกับวิทยานิพนธ์และดุษฎีนิพนธ์

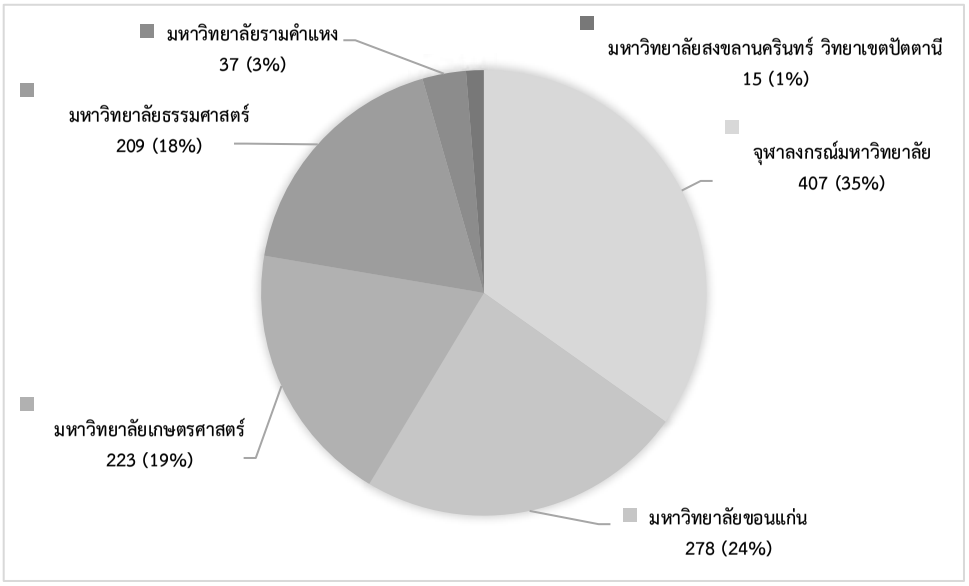
จากการสืบค้นข้อมูลเบื้องต้น ที่มีวิทยานิพนธ์และดุษฎีนิพนธ์สาขาสังคมวิทยาที่แล้วเสร็จและทำเป็นตัวเล่มเอกสาร ระหว่างปี พ.ศ. 2512-2560 จำนวน 1,169 เรื่อง โดยมีรายละเอียดดังแผนภูมิที่ 4.1



แผนภูมิที่ 4.1 แสดงจำนวนของวิทยานิพนธ์และดุษฎีนิพนธ์ระหว่างปี พ.ศ. 2512-2560

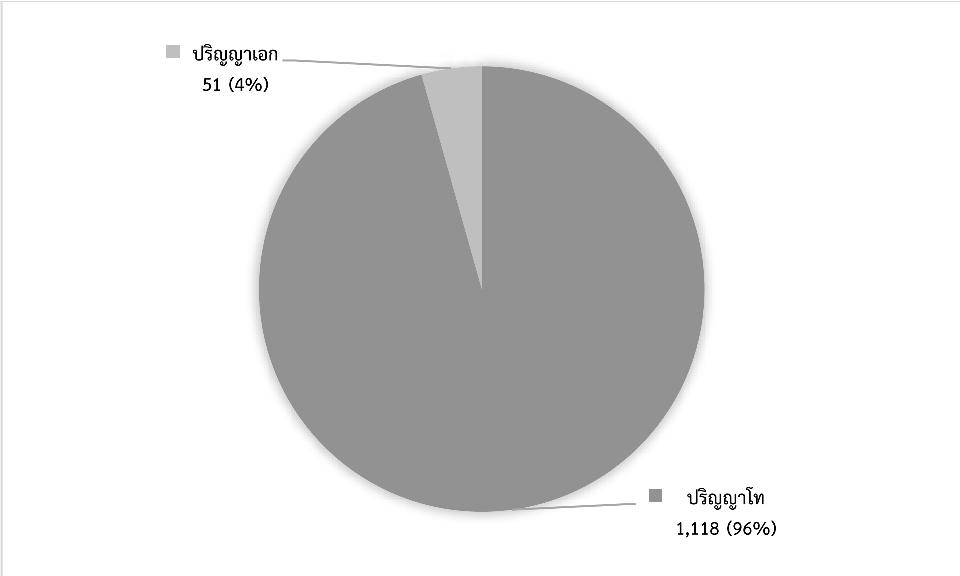
ตั้งแต่ปี พ.ศ. 2512 เป็นต้นมา มีจำนวนวิทยานิพนธ์และดุษฎีนิพนธ์เพิ่มขึ้นอย่างต่อเนื่อง จนถึงช่วงระหว่างปี พ.ศ. 2538 จำนวนวิทยานิพนธ์และดุษฎีนิพนธ์เริ่มมีปริมาณมากขึ้นอย่างชัดเจน โดยมีจำนวนมากที่สุดในปี พ.ศ. 2543 และหลังจากนั้นจำนวนวิทยานิพนธ์และดุษฎีนิพนธ์เริ่มลดลงอย่างต่อเนื่อง

มหาวิทยาลัยที่มีวิทยานิพนธ์และดุษฎีนิพนธ์สาขาสังคมวิทยามากที่สุดคือ จุฬาลงกรณ์มหาวิทยาลัย รองลงมาคือ มหาวิทยาลัยขอนแก่น มหาวิทยาลัยเกษตรศาสตร์ มหาวิทยาลัยธรรมศาสตร์ มหาวิทยาลัยรามคำแหง และมหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี ตามลำดับ โดยมีรายละเอียดดังแผนภูมิที่ 4.2



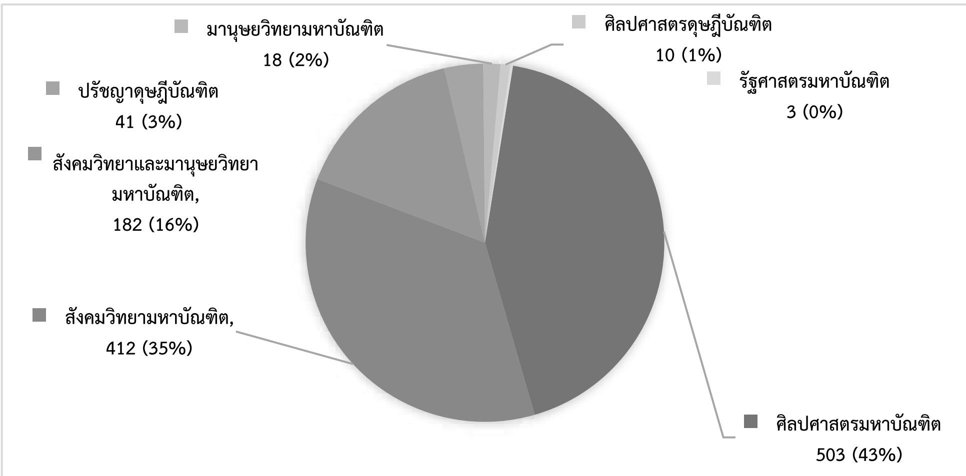
แผนภูมิที่ 4.2 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์ จำแนกตามมหาวิทยาลัย

หลักสูตรระดับสูงกว่าปริญญา คือ ปริญญาโท และปริญญาเอก ที่กำหนดให้ผู้เรียนทำวิทยานิพนธ์และดุษฎีนิพนธ์ แต่วิทยานิพนธ์และดุษฎีนิพนธ์ส่วนใหญ่เป็นวิทยานิพนธ์และดุษฎีนิพนธ์ระดับปริญญาโทมากกว่าระดับปริญญาเอก โดยมีรายละเอียดดังแผนภูมิที่ 4.3



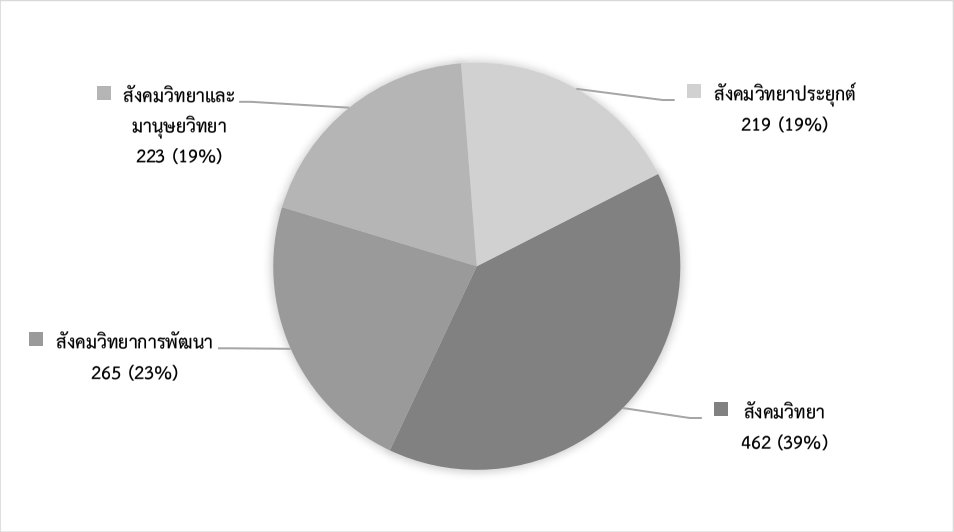
แผนภูมิที่ 4.3 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์ จำแนกตามระดับปริญญา

วิทยานิพนธ์และดุษฎีนิพนธ์เป็นปริญญาศิลปศาตรมหาบัณฑิตมากที่สุด รองลงมาคือ สังคมวิทยามหาบัณฑิต สังคมวิทยาและมานุษยวิทยามหาบัณฑิต ปรัชญาดุษฎีบัณฑิต มานุษยวิทยามหาบัณฑิต ศิลปศาสตรดุษฎีบัณฑิต และรัฐศาสตรมหาบัณฑิต (ปริญญาของแผนกสังคมวิทยา และภาคสังคมวิทยาและมนุษยวิทยา คณะรัฐศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย) โดยมีรายละเอียดดังแผนภูมิที่ 4.4



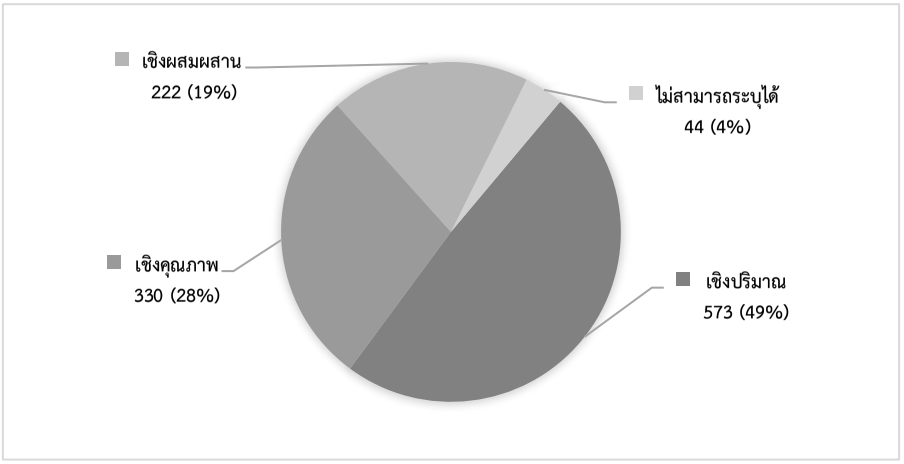
แผนภูมิที่ 4.4 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์ จำแนกตามชื่อปริญญา

หลักสูตรที่เกี่ยวข้องกับสาขาวิชาสังคมวิทยา มีการใช้ชื่อสาขาวิชาแตกต่างกันไม่มาก โดยสาขาที่มีวิทยานิพนธ์มากที่สุด คือ สังคมวิทยา รองลงมา คือ สังคมวิทยาการพัฒนา สังคมวิทยาและมานุษยวิทยา และสังคมวิทยาประยุกต์ ตามลำดับ โดยมีรายละเอียดดังแผนภูมิที่ 4.5



แผนภูมิที่ 4.5 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์ จำแนกตามชื่อสาขาวิชา

จากการจำแนกวิทยานิพนธ์และดุษฎีนิพนธ์สาขาสังคมวิทยาตั้งแต่ปี พ.ศ. 2512-2560 พบว่า เป็นงานวิจัยเชิงปริมาณมากที่สุด รองลงมา คือ เชิงคุณภาพ และเชิงผสมผสาน (ส่วนที่ไม่สามารถระบุได้ เนื่องจากไม่มีไฟล์ pdf ให้เปิดดูเพื่อวิเคราะห์และจำแนกว่าเป็นงานวิจัยประเภทใด) โดยมีรายละเอียดดังแผนภูมิที่ 4.6



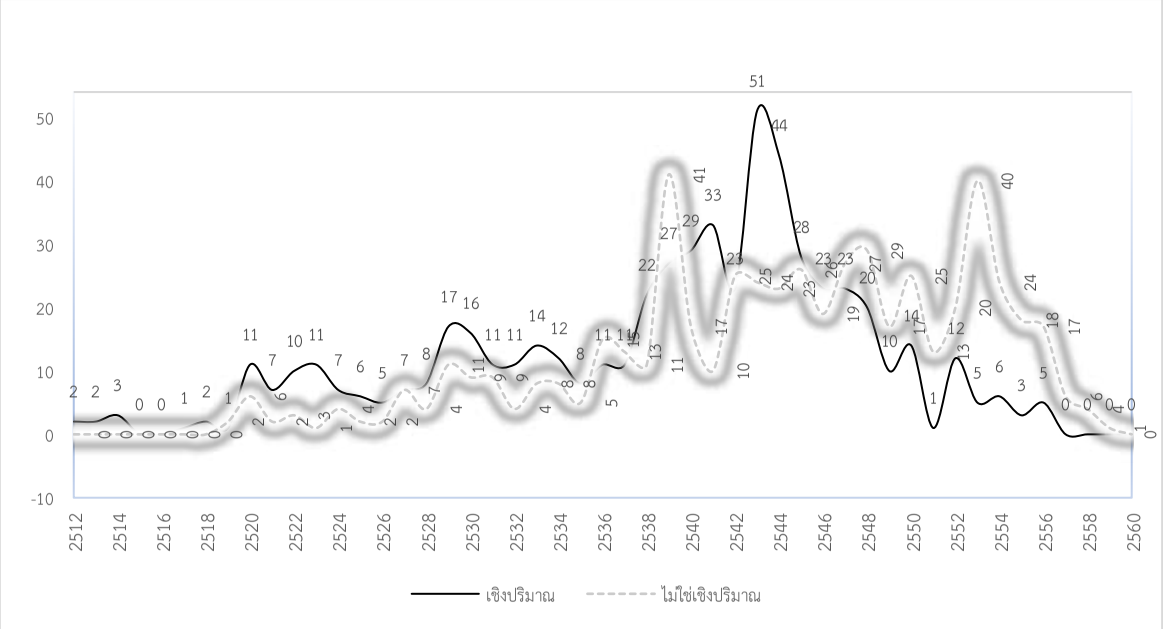
แผนภูมิที่ 4.6 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์ จำแนกตามวิธีการวิจัย

ข้อมูลทั่วไปเกี่ยวกับวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณ

เนื่องจากมีวิทยานิพนธ์และดุษฎีนิพนธ์ที่ไม่สามารถระบุได้ว่าเป็นงานวิจัยประเภทใด เพราะไม่สามารถเข้าถึงแฟ้มข้อมูลของวิทยานิพนธ์และดุษฎีนิพนธ์ได้ จึงตัดออกจากชุดข้อมูลจำนวน 44 เรื่อง เหลือวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณ 573 เรื่อง และวิทยานิพนธ์และดุษฎีนิพนธ์ไม่ใช่เชิงปริมาณ 552 เรื่อง (เชิงคุณภาพ 330 เรื่อง และเชิงผสมผสาน 222 เรื่อง) รวมเป็น 1,125 เรื่อง

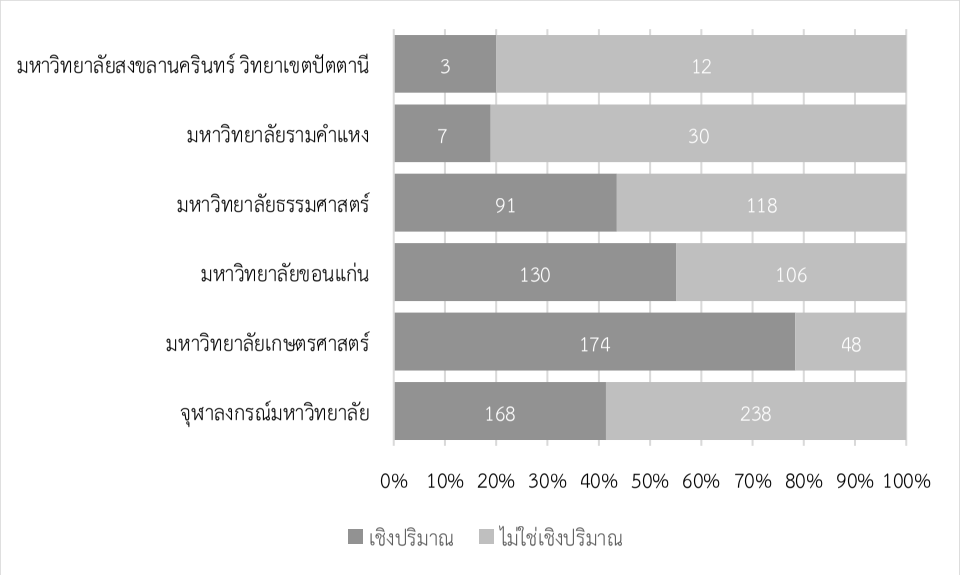
ช่วงแรกของการทำวิทยานิพนธ์และดุษฎีนิพนธ์สาขาสังคมวิทยาใช้วิธีเชิงปริมาณ จนถึงปี พ.ศ. 2519 วิทยานิพนธ์และดุษฎีนิพนธ์ที่ไม่ใช่เชิงปริมาณเริ่มปรากฏให้เห็นและเพิ่มขึ้นอย่างต่อเนื่อง โดยมีจำนวนมากที่สุดในปี พ.ศ. 2539 ซึ่งใกล้เคียงกับปี พ.ศ. 2553

เมื่อพิจารณาจากสัดส่วนโดยรวมพบว่า ส่วนใหญ่เป็นวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณมากกว่าไม่ใช่เชิงปริมาณ แต่มีจำนวนไม่แตกต่างกันมากนัก โดยมีรายละเอียดดังแผนภูมิที่ 4.7



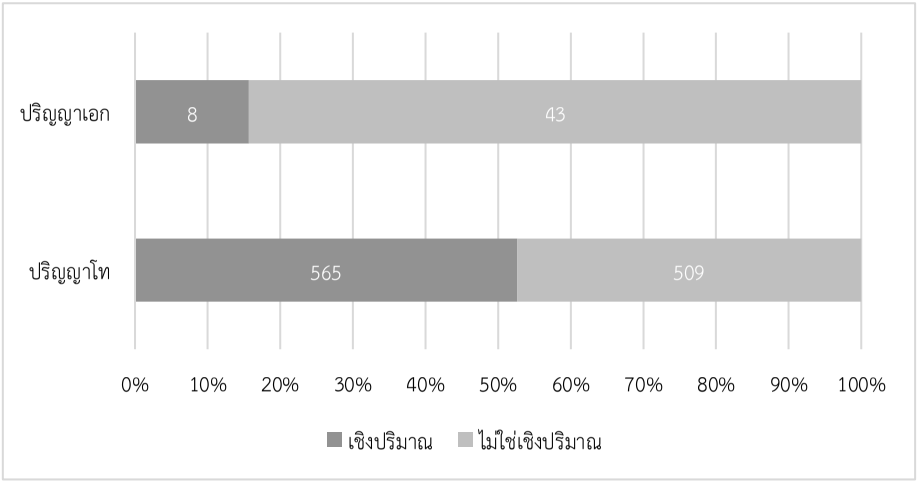
แผนภูมิที่ 4.7 แสดงจำนวนของวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณระหว่างปี พ.ศ. 2512-2560

เมื่อจำแนกตามมหาวิทยาลัย พบว่า มหาวิทยาลัยที่มีวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณมากกว่าไม่ใช่เชิงปริมาณ คือ มหาวิทยาลัยเกษตรศาสตร์และมหาวิทยาลัยขอนแก่น ส่วนมหาวิทยาลัยอื่นๆ มีจำนวนวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณน้อยกว่าไม่ใช่เชิงปริมาณ โดยมีรายละเอียดดังแผนภูมิที่ 4.8



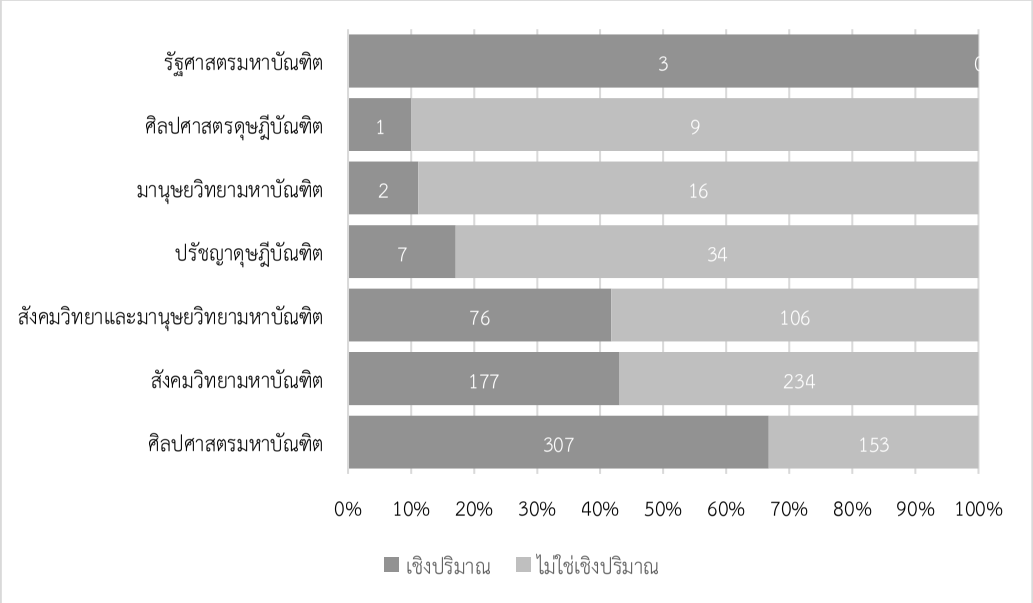
แผนภูมิที่ 4.8 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณและไม่ใช่เชิงปริมาณ จำแนกตามมหาวิทยาลัย

เมื่อจำแนกตามระดับปริญญา พบว่า วิทยานิพนธ์และดุษฎีนิพนธ์ระดับปริญญาโท เป็นวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณมากกว่าไม่ใช่เชิงปริมาณ ส่วนในระดับปริญญาเอกมีวิทยานิพนธ์และดุษฎีนิพนธ์ที่เป็นเชิงปริมาณน้อยกว่าไม่ใช่เชิงปริมาณ โดยมีรายละเอียดดังแผนภูมิที่ 4.9



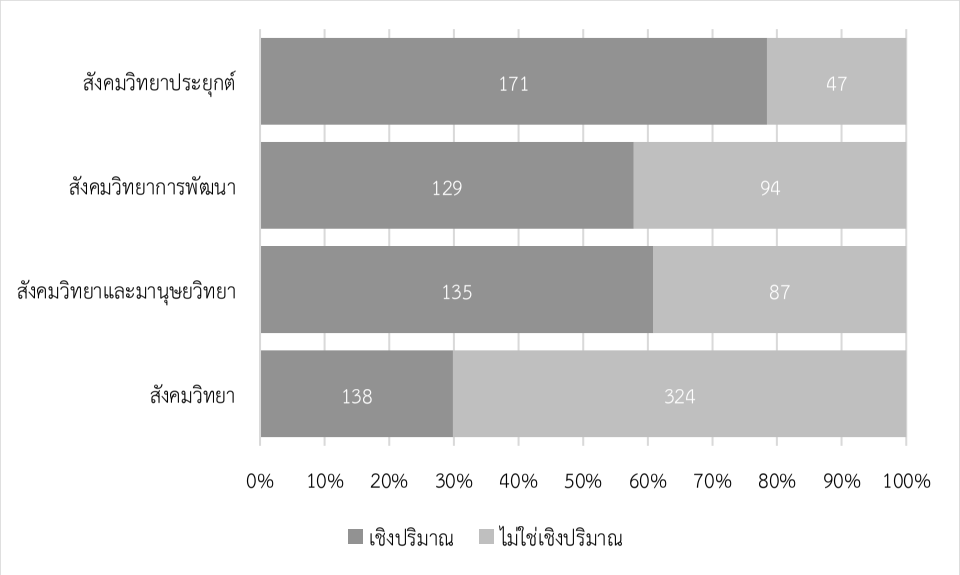
แผนภูมิที่ 4.9 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณและไม่ใช่เชิงปริมาณ จำแนกตามระดับปริญญา

ปริญญาที่มีการทำวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณมากกว่าไม่ใช่เชิงปริมาณ คือ ศิลปศาสตรมหาบัณฑิตและรัฐศาสตรมหาบัณฑิต ส่วนปริญญาอื่นๆ เป็นวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณน้อยกว่าไม่ใช่เชิงปริมาณ โดยมีรายละเอียดดังแผนภูมิที่ 4.10



แผนภูมิที่ 4.10 แสดงจำนวนและร้อยละของวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณและไม่ใช่เชิงปริมาณ จำแนกตามชื่อปริญญา

ส่วนใหญ่เกือบทุกสาขาวิชาเป็นวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณมากกว่าไม่ใช่เชิงปริมาณ คือ สังคมวิทยาประยุกต์ สังคมวิทยาและมานุษยวิทยา และสังคมวิทยาพัฒนาการ ยกเว้นสาขาวิชาสังคมวิทยา ที่เป็นวิทยานิพนธ์และดุษฎีนิพนธ์ไม่ใช่เชิงปริมาณมากกว่าเชิงปริมาณ โดยมีรายละเอียดดังแผนภูมิที่ 4.11



แผนภูมิที่ 4.11 แสดงจำนวนและร้อยละวิทยานิพนธ์และดุษฎีนิพนธ์เชิงปริมาณและไม่ใช่เชิงปริมาณ จำแนกตามชื่อสาขาวิชา

บทที่ 5
ผลการวิจัย

เนื้อหาผลการวิจัยแบ่งออกเป็น 3 ส่วน คือ ข้อมูลทั่วไปของกลุ่มตัวอย่าง การวิเคราะห์ความเสี่ยงของการเกิดความคาดเคลื่อน และการวิเคราะห์การทำนายการเกิดความเสี่ยง

ข้อมูลทั่วไปของกลุ่มตัวอย่าง

การรวบรวมข้อมูลจากขนาดตัวอย่างที่คำนวณได้ 236 ตัวอย่าง (เรื่อง) กลุ่มตัวอย่างมีลักษณะข้อมูลดังรายละเอียดต่อไปนี้

งานวิจัย (วิทยานิพนธ์และดุษฎีนิพนธ์) ที่เป็นกลุ่มตัวอย่างมากที่สุด คือ สาขาวิชาสังคมวิทยาประยุกต์ รองลงมา คือ สาขาสังคมวิทยา สาขาสังคมวิทยาและมานุษยวิทยา และน้อยที่สุด คือ สาขาสังคมวิทยาการพัฒนา ดังรายละเอียดในตารางที่ 5.1

ตารางที่ 5.1 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามสาขาวิชา

สาขาวิชา	จำนวน (เรื่อง)	รวม (ร้อยละ)
สังคมวิทยาประยุกต์	66	27.97
สังคมวิทยา	60	25.42
สังคมวิทยาการพัฒนา	56	23.73
สังคมวิทยาและมานุษยวิทยา	54	22.88
รวม	236	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างเป็นงานวิจัยระดับปริญญาโทมากกว่าระดับปริญญาเอก ดังรายละเอียดในตารางที่ 5.2

ตารางที่ 5.2 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามระดับปริญญา

ระดับปริญญา	จำนวน (เรื่อง)	รวม (ร้อยละ)
ปริญญาโท	232	98.31
ปริญญาเอก	4	1.69
รวม	236	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างมีการตั้งสมมติฐาน (มีการทดสอบสมมติฐาน) มากกว่าไม่มีการตั้งสมมติฐาน (ไม่มีการทดสอบสมติฐาน) ดังรายละเอียดในตารางที่ 5.3

ตารางที่ 5.3 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามการตั้งสมมติฐาน

สมมติฐาน	จำนวน (เรื่อง)	รวม (ร้อยละ)
มี	225	95.34
ไม่มี	11	4.66
รวม	236	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างมีการตรวจสอบเครื่องมือมากกว่าไม่มีการตรวจสอบเครื่องมือ ดังรายละเอียดในตารางที่ 5.4

ตารางที่ 5.4 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามการตรวจสอบเครื่องมือ

การตรวจสอบเครื่องมือ	จำนวน (เรื่อง)	รวม (ร้อยละ)
มี	167	70.76
ไม่มี	69	29.24
รวม	236	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างสามารถจำแนกได้เป็น 3 กลุ่ม โดยพบว่า มีตรวจสอบความถูกต้องและความเที่ยงตรงมากที่สุด รองลงมา คือ ตรวจสอบความเที่ยงตรง และความถูกต้องตามลำดับ ดังรายละเอียดในตารางที่ 5.5

ตารางที่ 5.5 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามวิธีการตรวจสอบเครื่องมือ

วิธีการตรวจสอบเครื่องมือ	จำนวน (เรื่อง)	รวม (ร้อยละ)
ความถูกต้องและความเที่ยงตรง	140	59.32
ความเที่ยงตรง	15	6.36
ความถูกต้อง	12	5.08
ไม่มีการตรวจสอบ	69	29.24
รวม	236	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างใช้ระดับการวัดตัวแปรเป็นแบบค่าแบ่งกลุ่ม (นามมาตราและอันดับมาตรา) มากกว่าค่าต่อเนื่อง (ช่วงมาตราและอัตราส่วนมาตรา) ดังรายละเอียดในตารางที่ 5.6

ตารางที่ 5.6 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามระดับการวัดของตัวแปร

ระดับการวัดของตัวแปร	จำนวน (เรื่อง)	รวม (ร้อยละ)
ตัวแปรค่าแบ่งกลุ่ม	130	55.08
ตัวแปรค่าต่อเนื่อง	106	44.92
รวม	236	100.00

เมื่อพิจารณาในภาพรวมพบว่า เป็นตัวแปรอันดับมาตรามากที่สุด รองลงมา คือ ตัวแปรช่วงมาตรา ตัวแปรนามมาตรา และน้อยที่สุด คือ ตัวแปรอัตราส่วนมาตรา

หากจำแนกเป็นระดับการวัดของตัวแปร พบว่า ตัวแปรค่าแบ่งกลุ่มเป็นตัวแปรอันดับมาตรามากกว่าตัวแปรนามมาตรา และตัวแปรค่าต่อเนื่องเป็นตัวแปรช่วงมาตรามากกว่าตัวแปรอันตราส่วนมาตรา ดังรายละเอียดในตารางที่ 5.7

ตารางที่ 5.7 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามประเภทของตัวแปร

ประเภทของตัวแปร	จำนวน (เรื่อง)	รวม (ร้อยละ)
ตัวแปรค่าแบ่งกลุ่ม		
- ตัวแปรอันดับมาตรา	136	53.33
- ตัวแปรนามมาตรา	12	4.71
ตัวแปรค่าต่อเนื่อง		
- ตัวแปรช่วงมาตรา	100	39.22
- ตัวแปรอัตราส่วนมาตรา	7	2.75
รวม (คำนวณแบบหลายคำตอบ)	255	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่าง ประชากรที่นำมาใช้ในการศึกษาเป็นประชากรที่ทราบจำนวนมากกว่าไม่ทราบจำนวน ดังรายละเอียดในตารางที่ 5.8

ตารางที่ 5.8 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามประเภทของประชากร

ประเภทของประชากร	จำนวน (เรื่อง)	รวม (ร้อยละ)
ทราบจำนวน	176	74.58
ไม่ทราบจำนวน	60	25.42
รวม	236	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างใช้วิธีการกำหนดขนาดตัวอย่างแบบไม่ใช้วิธีการทางสถิติมากกว่าใช้วิธีการทางสถิติ ดังรายละเอียดในตารางที่ 5.9

ตารางที่ 5.9 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามวิธีการกำหนดขนาดตัวอย่าง

การกำหนดขนาดตัวอย่าง	จำนวน (เรื่อง)	รวม (ร้อยละ)
ไม่ใช้วิธีการทางสถิติ	129	54.66
ใช้วิธีการทางสถิติ	107	45.34
รวม	236	100.0

งานวิจัยที่เป็นกลุ่มตัวอย่างใช้วิธีการเลือกตัวอย่างแบบใช้หลักความน่าจะเป็นมากกว่าไม่ใช้หลักความน่าจะเป็น ดังรายละเอียดในตารางที่ 5.10

ตารางที่ 5.10 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามประเภทการเลือกตัวอย่าง

ประเภทการเลือกตัวอย่าง	จำนวน (เรื่อง)	รวม (ร้อยละ)
ใช้หลักความน่าจะเป็น	138	58.47
ไม่ใช้หลักความน่าจะเป็น	98	41.53
รวม	236	100.00

วิธีการเลือกตัวอย่างของงานวิจัยที่เป็นกลุ่มตัวอย่าง เมื่อพิจารณาในภาพรวมพบว่า ใช้วิธีการเลือกตัวอย่างแบบสุ่มอย่างง่ายมากที่สุด รองลงมา คือ การเลือกตัวอย่างแบบเจาะจง การเลือกตัวอย่างแบบสุ่มด้วยการแบ่งชั้น/ชั้นภูมิ การเลือกตัวอย่างแบบสุ่มอย่างเป็นระบบ การเลือกตัวอย่างแบบตามสะดวก/เผอิญพบ การเลือกตัวอย่างแบบสัดส่วนหรือโควต้า และน้อยที่สุด คือ การเลือกตัวอย่างแบบสุ่มเชิงพื้นที่

หากจำแนกเป็นประเภทการเลือกตัวอย่าง พบว่า การเลือกตัวอย่างแบบใช้หลักความน่าจะเป็น ใช้วิธีการเลือกตัวอย่างแบบสุ่มอย่างง่ายมากที่สุด รองลงมา คือ การเลือกตัวอย่างแบบสุ่มด้วยการแบ่งชั้น/ชั้นภูมิ การเลือกตัวอย่างแบบสุ่มอย่างเป็นระบบ และน้อยที่สุด คือ การเลือกตัวอย่างแบบสุ่มเชิงพื้นที่ ส่วนการเลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น ใช้การเลือกตัวอย่างแบบเจาะจงมากที่สุด รองลงมา คือ การเลือกตัวอย่างแบบตามสะดวก/เผอิญพบ และน้อยที่สุด คือ การเลือกตัวอย่างแบบสัดส่วนหรือโควตา ดังรายละเอียดในตารางที่ 5.11

ตารางที่ 5.11 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามวิธีการเลือกตัวอย่าง

วิธีการเลือกตัวอย่าง	จำนวน (เรื่อง)	รวม (ร้อยละ)
การเลือกตัวอย่างแบบใช้หลักความน่าจะเป็น		
- การเลือกตัวอย่างแบบสุ่มอย่างง่าย	84	29.68
- การเลือกตัวอย่างแบบสุ่มด้วยการแบ่งชั้น/ชั้นภูมิ	59	20.85
- การเลือกตัวอย่างแบบสุ่มอย่างเป็นระบบ	29	10.25
- การเลือกตัวอย่างแบบสุ่มเชิงพื้นที่	3	1.06
การเลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น		
- การเลือกตัวอย่างแบบเจาะจง	81	28.62
- การเลือกตัวอย่างแบบตามสะดวก/เผอิญพบ	19	6.71
- การเลือกตัวอย่างแบบสัดส่วนหรือโควตา	8	2.83
รวม (คำนวณแบบหลายคำตอบ)	283	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างเกือบทั้งหมดไม่มีการตรวจสอบข้อมูลก่อนการวิเคราะห์ข้อมูลด้วยสถิติ ดังรายละเอียดในตารางที่ 5.12

ตารางที่ 5.12 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามการตรวจสอบข้อมูลเบื้องต้น

การตรวจสอบข้อมูลเบื้องต้น	จำนวน (เรื่อง)	รวม (ร้อยละ)
ไม่ตรวจสอบ	234	99.15
ตรวจสอบ	2	0.85
รวม	236	100.00

วิธีการตรวจสอบข้อมูลที่พบเป็นการตรวจสอบข้อมูลตามข้อตกลงเบื้องต้นของสถิติกลุ่มแบบจำลองถดถอยแบบพหุ คือ การตรวจสอบค่าสูญหาย การตรวจสอบความเท่ากันของความแปรปรวนของความคลาดเคลื่อน การตรวจสอบความเป็นเส้นตรง การตรวจสอบภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรอิสระหลายตัวแปร การตรวจสอบความสัมพันธ์ภายในตัวแปร การตรวจสอบความสุ่ม และการตรวจสอบค่าส่วนเหลือ ดังรายละเอียดในตารางที่ 5.13

ตารางที่ 5.13 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามวิธีการตรวจสอบข้อมูลเบื้องต้น

วิธีการตรวจสอบข้อมูล	จำนวน (เรื่อง)	รวม (ร้อยละ)
การตรวจสอบค่าสูญหาย (missing)	1	10.00
การตรวจสอบความเท่ากันของความแปรปรวนของความคลาดเคลื่อน (homoscedasticity)	2	20.00
การตรวจสอบความเป็นเส้นตรง (linearity)	2	20.00
การตรวจสอบภาวะร่วมเชิงเส้นตรงระหว่างตัว แปรอิสระหลายตัวแปร (multicollinearity)	2	20.00
การตรวจสอบความสัมพันธ์ภายในตัวแปร (autocorrelation)	1	10.00
การตรวจสอบความสุ่ม (randomness)	1	10.00
การตรวจสอบค่าส่วนเหลือ (residuals)	1	10.00
รวม (คำนวณแบบหลายคำตอบ)	10	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างที่มีการวิเคราะห์ข้อมูลเบื้องต้นก่อนการวิเคราะห์ข้อมูลด้วยสถิติ 2 เรื่อง พบว่า มีงานวิจัย 1 เรื่อง ที่มีการตรวจสอบข้อมูลตามข้อตกลงเบื้องต้นและมีการปรับแก้ข้อมูลด้วย ดังรายละเอียดในตารางที่ 5.14

ตารางที่ 5.14 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามการปรับแก้ข้อมูล

การปรับแก้ข้อมูล	จำนวน (เรื่อง)	รวม (ร้อยละ)
ไม่ปรับแก้	235	99.58
ปรับแก้	1	0.42
รวม	236	100.00

งานวิจัยที่เป็นกลุ่มตัวอย่างมีการใช้สถิติเชิงอ้างอิงมากกว่าสถิติเชิงพรรณนา โดยสถิติเชิงอ้างอิงที่ใช้เป็นสถิติแบบมีพารามิเตอร์มากกว่าสถิติแบบไม่มีพารามิเตอร์ ดังรายละเอียดในตารางที่ 5.15

ตารางที่ 5.15 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามรูปแบบการใช้สถิติ

รูปแบบการใช้สถิติ	จำนวน (เรื่อง)	รวม (ร้อยละ)
สถิติเชิงพรรณนา	31	13.14
สถิติเชิงอ้างอิง		
- สถิติแบบไม่มีพารามิเตอร์	100	42.37
- สถิติแบบมีพารามิเตอร์	105	44.49
รวม	236	100.00

ประเภทการใช้สถิติในการวิเคราะห์ข้อมูลของงานวิจัยที่เป็นกลุ่มตัวอย่าง พบว่า ใช้การวิเคราะห์ความเป็นอิสระของข้อมูลด้วยไคสแควร์มากที่สุด รองลงมา คือ การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน การทดสอบสอบค่าเฉลี่ยระหว่าง 2 กลุ่มตัวอย่างด้วยค่าที ค่าร้อยละ การวิเคราะห์ความแปรปรวนแบบทางเดียว ค่าเฉลี่ย การวิเคราะห์ความแปรปรวนแบบสองทางและการวิเคราะห์เส้นทางเท่ากัน การวิเคราะห์การถดถอยเชิงโลจิสติกพหุ และน้อยที่สุด คือ การทดสอบค่าเฉลี่ยระหว่างกลุ่มตัวอย่างที่มีความสัมพันธ์กันด้วยค่าทีและการวิเคราะห์จำแนก ดังรายละเอียดในตารางที่ 5.16

ตารางที่ 5.16 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามประเภทสถิติ

ประเภทสถิติ	จำนวน (เรื่อง)	รวม (ร้อยละ)
สถิติพรรณนา		
- ค่าร้อยละ (percentage)	31	9.09
- ค่าเฉลี่ย (mean)	14	4.11
สถิติเชิงอ้างอิงแบบมีพารามิเตอร์		
- การทดสอบสอบค่าเฉลี่ยระหว่าง 2 กลุ่มตัวอย่างด้วยค่าที (two-sample t-test)	42	12.32
- การทดสอบค่าเฉลี่ยระหว่างกลุ่มตัวอย่างที่มีความสัมพันธ์กันด้วยค่าที (paired-sample t-test)	1	0.29
- การวิเคราะห์ความแปรปรวนแบบทางเดียว (one-way ANOVA)	30	8.80
- การวิเคราะห์ความแปรปรวนแบบสองทาง (two-way ANOVA)	4	1.17
- การวิเคราะห์สหสัมพันธ์แบบเพียร์สัน (pearson correlation)	47	13.78
- การวิเคราะห์การถดถอยเชิงเส้นตรงแบบพหุ (multiple regression)	48	14.08
- การวิเคราะห์เส้นทาง (path analysis)	4	1.17
- การวิเคราะห์จำแนก (classification analysis)	1	0.29
สถิติเชิงอ้างอิงแบบไม่มีพารามิเตอร์		
- การวิเคราะห์ความเป็นอิสระของข้อมูลด้วยไคสแควร์ (chi-square)	117	34.49
- การวิเคราะห์การถดถอยเชิงโลจิสติกพหุ (multiple logistic regression)	2	0.59
รวม (คำนวณแบบหลายคำตอบ)	341	100.00

การวิเคราะห์ความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย

การวิเคราะห์ความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย เป็นการพิจารณาจากตัวแปรที่เป็นองค์ประกอบภายในกรอบแนวคิดการวิจัยจำนวน 8 ตัวแปร คือ สมมติฐาน สถิติ เครื่องมือวัด ระดับการวัด การตรวจสอบข้อมูล ประชากร ขนาดตัวอย่าง และการเลือกตัวอย่าง โดยมีเงื่อนไขในการพิจารณาแต่ละตัวแปรตามตารางที่ 3.2 แสดงเงื่อนไขของตัวแปรหุ่นที่มีผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย และเพิ่มเงื่อนไขตัวแปรสมมติฐานและสถิติ ที่มีและไม่มีความเสี่ยงของการเกิดความคลาดคลื่อนในการวิจัย ดังต่อไปนี้

1. สมมติฐาน งานวิจัยที่มีสมมติฐาน = มีความเสี่ยงของการเกิดความคลาดคลื่อนในการวิจัย และงานวิจัยที่ไม่มีสมมติฐาน = ไม่มีความเสี่ยงของการเกิดความคลาดคลื่อนในการวิจัย

2. สถิติ งานวิจัยที่ใช้สถิติแบบมีพารามิเตอร์ = มีความเสี่ยงของการเกิดความคลาดคลื่อนในการวิจัย และงานวิจัยที่ใช้สถิติแบบไม่มีพารามิเตอร์และสถิติเชิงพรรณนา = ไม่มีความเสี่ยงของการเกิดความคลาดคลื่อนในการวิจัย

จากผลการวิเคราะห์พบว่า ตัวแปรที่น่าจะมีผลทำให้เกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากที่สุด คือ การตรวจสอบข้อมูล รองลงมา คือ สมมติฐาน ขนาดตัวอย่าง ระดับการวัด สถิติ การเลือกตัวอย่าง เครื่องมือวัด และน้อยที่สุด คือ ประชากร โดยมีรายละเอียดดังในตารางที่ 5.17

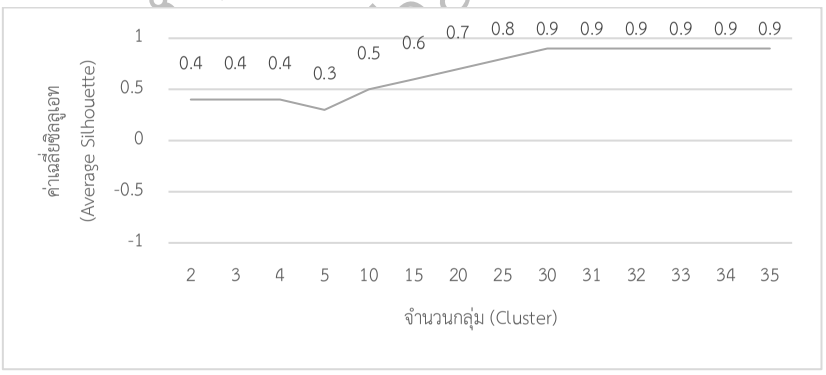
ตารางที่ 5.17 แสดงจำนวนและร้อยละกลุ่มตัวอย่าง จำแนกตามความเสี่ยงของการเกิดความคลาดเคลื่อน

ตัวแปรหุ่น	ความเสี่ยง		รวม
	มี	ไม่มี	
สมมติฐาน	95.34 (225)	4.66 (11)	100.00 (236)
สถิติ	44.49 (105)	55.51 (131)	100.00 (236)
เครื่องมือวัด	29.24 (69)	70.76 (167)	100.00 (236)
ระดับการวัด	44.92 (106)	55.08 (130)	100.00 (236)
การตรวจสอบข้อมูล	99.15 (234)	0.85 (2)	100.00 (236)
ประชากร	25.42 (60)	74.58 (176)	100.00 (236)
ขนาดตัวอย่าง	54.66 (129)	45.34 (107)	100.00 (236)
การเลือกตัวอย่าง	41.53 (98)	58.47 (138)	100.00 (236)

การจัดกลุ่มความเสี่ยงการเกิดความคลาดเคลื่อนมีตัวแปรหุ่น 8 ตัว คือ สมมติฐาน สถิติ เครื่องมือวัด ระดับการวัด การตรวจสอบข้อมูล ประชากร ขนาดตัวอย่าง และการเลือกตัวอย่าง โดยใช้การวิเคราะห์จัดกลุ่ม (Cluster Analysis: CA) เพื่อหากลุ่มตัวอย่างที่มีความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย โดยมีรายละเอียดดังต่อไปนี้

การวิเคราะห์หาจำนวนกลุ่ม (clusters) ที่เหมาะสมด้วยวิธีการวิเคราะห์จัดกลุ่มแบบสองขั้นตอน (Two Step Cluster Analysis: TSCA) ได้จำนวนกลุ่มตามที่โปรแกรมกำหนด (determine automatically) จำนวน 2 กลุ่ม แต่เมื่อพิจารณาจากคุณภาพแบบจำลองโดยรวมของกลุ่ม (cluster quality) ได้ค่าเฉลี่ยซิลลูเอท (Average Silhouette) = 0.4 แสดงว่า การจัดกลุ่มมีการรวมกลุ่ม (cohesion) และแบ่งแยก (separation) อยู่ในเกณฑ์ปานกลางและยังไม่ดีพอ (ค่าเฉลี่ยซิลลูเอทมีค่าระหว่าง -1 ถึง 1)

เมื่อนำไปทดลองจัดกลุ่มแบบเคมีน (K-mean Cluster Analysis) ยังไม่สามารถแสดงวิทยานิพนธ์หรือดุษฎีนิพนธ์ที่มีการตรวจสอบคุณภาพข้อมูลออกมาจำนวน 2 เรื่อง จึงทำการกำหนดจำนวนกลุ่มแบบกำหนดเอง (specify fixed) โดยเพิ่มจำนวนกลุ่มครั้งละ 5 กลุ่มจนถึง 30 กลุ่ม จึงได้ค่าเฉลี่ยซิลลูเอท = 0.9 แสดงว่าข้อมูลภายในกลุ่มเดียวกันมีความคล้ายกันดีมาก และข้อมูลระหว่างกลุ่มมีความแตกต่างกันดีมาก และเมื่อเพิ่มจำนวนกลุ่มครั้งละ 1 กลุ่มจนถึง 35 กลุ่ม ยังได้ค่าเฉลี่ยซิลลูเอทคงที่ไม่มีการเปลี่ยนแปลง โดยมีรายละเอียดดังในภาพที่ 4.1



แผนภูมิที่ 5.1 แสดงคุณภาพการจัดกลุ่มด้วยวิธีสองขั้นตอน

จากภาพที่ 4.1 แสดงว่าจำนวนกลุ่มที่เหมาะสมและอยู่ในเกณฑ์ดีและมีความคงที่ คือ 30 กลุ่ม จึงนำไปวิเคราะห์จัดกลุ่มแบบไม่เป็นขั้นตอน (Nonhierarchical Cluster Analysis: NCA) หรือการวิเคราะห์จัดกลุ่มแบบเคมีน (K-means Cluster Analysis)

ผลการวิเคราะห์จัดกลุ่มแบบเคมีน ได้รูปแบบการวิธีการดำเนินการวิจัย 30 กลุ่ม ดังรายละเอียดในตารางที่ 5.18

ตารางที่ 5.18 แสดงจำนวนการจัดกลุ่มรูปแบบวิธีการดำเนินการวิจัยแต่ละรูปแบบของกลุ่มตัวอย่าง

กลุ่ม	สมมติฐาน	เครื่องมือวัด	ประชากร	ขนาดตัวอย่าง	การเลือกตัวอย่าง	ระดับการวัด	การตรวจสอบข้อมูล	สถิติ	จำนวนกลุ่มตัวอย่าง (เรื่อง)
1	0	0	1	0	1	0	0	0	6
2	1	0	1	0	1	0	0	0	24
3	1	0	1	0	0	0	0	1	3
4	1	1	1	1	1	0	0	0	53
5	1	0	0	0	1	0	0	1	2
6	1	0	1	0	1	0	0	1	7
7	1	0	1	0	0	1	0	1	10
8	1	1	1	0	0	1	0	1	14
9	1	0	1	1	1	1	0	1	3
10	1	1	1	0	0	0	0	0	18
11	1	1	1	0	1	0	0	1	5
12	1	1	0	1	1	1	0	0	2
13	1	0	1	0	0	0	0	0	7
14	1	1	1	1	0	1	0	0	7
15	1	0	0	0	0	1	0	1	2
16	0	0	0	0	0	0	0	0	1
17	1	1	1	0	0	1	1	1	2
18	1	1	1	0	1	1	0	0	6
19	1	1	0	0	0	0	0	0	6
20	1	0	0	0	1	0	0	0	1
21	0	1	1	1	0	0	0	0	1
22	1	1	1	0	0	0	0	1	6
23	0	1	1	0	0	1	0	0	3
24	1	0	0	0	0	0	0	1	1
25	1	0	0	0	1	1	0	0	1
26	1	1	1	0	0	1	0	0	12
27	1	0	0	0	0	0	0	0	2
28	1	1	1	1	1	1	0	1	22
29	1	0	1	1	1	0	0	1	4
30	1	1	1	1	0	0	0	1	5

จากตารางที่ 5.18 มีกลุ่มตัวอย่างที่ไม่มีการทดสอบสมมติฐาน 4 กลุ่ม จำนวน 11 เรื่อง คือ กลุ่มที่ 1 จำนวน 6 เรื่อง กลุ่มที่ 16 จำนวน 1 เรื่อง กลุ่มที่ 21 จำนวน 1 เรื่อง และกลุ่มที่ 23 จำนวน 3 เรื่อง จึงตัดกลุ่มตัวอย่างชุดนี้ออกจากการวิเคราะห์ความเสี่ยง

จากการตัดกลุ่มตัวอย่างที่ไม่มีการทดสอบสมมติฐานออก จึงเหลือกลุ่มตัวอย่างที่มีการทดสอบสมมติฐาน 26 กลุ่ม จำนวน 225 เรื่อง และนำมากำหนดความเสี่ยง โดยให้ค่า 1 = มีความเสี่ยง 0 = ไม่มีความเสี่ยง ดังมีรายละเอียดในตารางที่ 5.19

ตารางที่ 5.19 แสดงจำนวนกลุ่มตัวอย่างที่มีการทดสอบสมมติฐาน จำแนกตามความเสี่ยง

กลุ่ม	สมมติฐาน	เครื่องมือวัด	ประชากร	ขนาดตัวอย่าง	การเลือกตัวอย่าง	ระดับการวัด	การตรวจสอบข้อมูล	สถิติ	ความเสี่ยง	จำนวนกลุ่มตัวอย่าง (เรื่อง)
2	1	0	1	0	1	0	0	0	1	24
3	1	0	1	0	0	0	0	1	1	3
4	1	1	1	1	1	0	0	0	0	53
5	1	0	0	0	1	0	0	1	1	2
6	1	0	1	0	1	0	0	1	1	7
7	1	0	1	0	0	1	0	1	1	10
8	1	1	1	0	0	1	0	1	1	14
9	1	0	1	1	1	1	0	1	1	3
10	1	1	1	0	0	0	0	0	0	18
11	1	1	1	0	1	0	0	1	1	5
12	1	1	0	1	1	1	0	0	0	2
13	1	0	1	0	0	0	0	0	1	7
14	1	1	1	1	0	1	0	0	0	7
15	1	0	0	0	0	1	0	1	1	2
17	1	1	1	0	0	1	1	1	1	2
18	1	1	1	0	1	1	0	0	0	6
19	1	1	0	0	0	0	0	0	0	6
20	1	0	0	0	1	0	0	0	1	1
22	1	1	1	0	0	0	0	1	1	6
24	1	0	0	0	0	0	0	1	1	1
25	1	0	0	0	1	1	0	0	1	1
26	1	1	1	0	0	1	0	0	0	12
27	1	0	0	0	0	0	0	0	1	2
28	1	1	1	1	1	1	0	1	1	22
29	1	0	1	1	1	0	0	1	1	4
30	1	1	1	1	0	0	0	1	1	5

จากตารางที่ 5.19 แสดงให้เห็นว่ามีกลุ่มตัวอย่างที่มีความเสี่ยงเกิดความคลาดเคลื่อนมีจำนวน 19 กลุ่ม คือ กลุ่ม 2, 3, 5-9, 11, 13, 15, 17, 20, 22, 24-25 และ 27-30 และไม่มีความเสี่ยงหรือมีโอกาสน้อยในการเกิดความคลาดเคลื่อน 7 กลุ่ม คือ กลุ่ม 4, 10, 12, 14, 18, 19 และ 26

รูปแบบความเสี่ยงที่เกิดขึ้นแต่ละกลุ่มตัวอย่างมีสาเหตุที่แตกต่างกัน แต่เมื่อพิจารณาจากจำนวนกลุ่มตัวอย่างในกลุ่มที่มีความเสี่ยงเกิดความคลาดเคลื่อนมาก 5 อันดับ มีรายละเอียดดังในตารางที่ 5.20

ตารางที่ 5.20 แสดงจำนวนของกลุ่มตัวอย่าง จำแนกตามอันดับของความเสี่ยง

กลุ่ม	สมมติฐาน	เครื่องมือวัด	ประชากร	ขนาดตัวอย่าง	การเลือกตัวอย่าง	ระดับการวัด	การตรวจสอบข้อมูล	สถิติ	ความเสี่ยง	จำนวนกลุ่มตัวอย่าง (เรื่อง)
2	1	0	1	0	1	0	0	0	1	24
28	1	1	1	1	1	1	0	1	1	22
8	1	1	1	0	0	1	0	1	1	14
7	1	0	1	0	0	1	0	1	1	10
6	1	0	1	0	1	0	0	1	1	7
13	1	0	1	0	0	0	0	0	1	7
22	1	1	1	0	0	0	0	1	1	6
11	1	1	1	0	1	0	0	1	1	5
30	1	1	1	1	0	0	0	1	1	5
29	1	0	1	1	1	0	0	1	1	4
3	1	0	1	0	0	0	0	1	1	3
9	1	0	1	1	1	1	0	1	1	3
5	1	0	0	0	1	0	0	1	1	2
15	1	0	0	0	0	1	0	1	1	2
17	1	1	1	0	0	1	1	1	1	2
27	1	0	0	0	0	0	0	0	1	2
20	1	0	0	0	1	0	0	0	1	1
24	1	0	0	0	0	0	0	1	1	1
25	1	0	0	0	1	1	0	0	1	1

จากตารางที่ 5.20 พบว่า กลุ่มตัวอย่างที่มีความเสี่ยงเกิดความคลาดเคลื่อนมากที่สุดอันดับที่ 5 อันดับ มีทั้งกลุ่มตัวอย่างที่มีสมมติฐานและใช้สถิติเชิงอ้างอิงแบบมีพารามิเตอร์และไม่มีพารามิเตอร์ ดังนี้

1. อันดับที่ 1 คือ กลุ่มที่ 2 ใช้สถิติแบบไม่มีพารามิเตอร์ แม้ว่าจะเป็นสถิติที่เป็นอิสระจากข้อตกลง (assumption-free tests) แต่ไม่มีการตรวจสอบเครื่องมือวัด

2. อันดับที่ 2 คือ กลุ่มที่ 28 ใช้สถิติแบบมีพารามิเตอร์ โดยไม่มีการตรวจสอบข้อมูล

3. อันดับที่ 3 คือ กลุ่มที่ 8 ใช้สถิติแบบมีพารามิเตอร์ โดยไม่ใช้วิธีการทางสถิติในการกำหนดขนาดตัวอย่าง ระดับการวัดไม่เป็นค่าต่อเนื่อง และไม่มีการตรวจสอบข้อมูล

4. อันดับที่ 4 คือ กลุ่มที่ 7 ใช้สถิติแบบมีพารามิเตอร์ โดยไม่มีการตรวจสอบเครื่องมือวัด ไม่ใช้วิธีการทางสถิติในการกำหนดขนาดตัวอย่าง เลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น และไม่มีการตรวจสอบข้อมูล

5. อันดับที่ 5 คือ กลุ่มที่ 6 และกลุ่มที่ 13 โดยกลุ่มที่ 6 ใช้สถิติแบบมีพารามิเตอร์ โดยไม่มีการตรวจสอบเครื่องมือวัด ไม่ใช้วิธีการทางสถิติในการกำหนดขนาดตัวอย่าง ระดับการวัดไม่เป็นค่าต่อเนื่อง และไม่มีการตรวจสอบข้อมูล ส่วนกลุ่มที่ 13 แม้ว่าจะใช้สถิติแบบไม่มีพารามิเตอร์ แต่ไม่มีการตรวจสอบเครื่องมือ รวมถึงไม่ใช้วิธีการทางสถิติในการกำหนดขนาดตัวอย่าง เลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น ระดับการวัดไม่เป็นค่าต่อเนื่อง และไม่มีการตรวจสอบข้อมูล

เมื่อวิเคราะห์ภาพรวมโอกาสของการเกิดความเสี่ยง แต่ละกลุ่มความเสี่ยงมีจำนวนกลุ่มตัวอย่าง ดังมีรายละเอียดในตารางที่ 5.21

ตารางที่ 5.21 แสดงจำนวนและร้อยละของกลุ่มตัวอย่าง จำแนกตามความเสี่ยง

ความเสี่ยง	จำนวน (เรื่อง)	รวม (ร้อยละ)
มี	121	53.78
ไม่มี	104	46.22
รวม	225	100.00

จากตารางที่ 5.21 พบว่า มีวิทยานิพนธ์ที่เป็นกลุ่มตัวอย่างมีความเสี่ยงมากกว่าไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อน แต่มีจำนวนไม่แตกต่างกันมากนัก จึงทำการทดสอบสัดส่วนตามรายละเอียดในตารางที่ 5.19

ตารางที่ 5.22 แสดงการเปรียบเทียบจำนวนและสัดส่วนระหว่างค่าสังเกตกับค่าคาดหวังของความเสี่ยง

ความเสี่ยง	จำนวน (N)	สัดส่วนค่าสังเกต (Observed Prop.)	สัดส่วนการทดสอบ (Test Prop.)	Exact Sig.
มี	121	.54		
ไม่มี	104	.46	.50	.286
รวม	225	1.00		

จากตารางที่ 5.22 พบว่า งานวิจัยที่เป็นกลุ่มตัวอย่างมีสัดส่วนระหว่างมีความเสี่ยงและไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อนไม่แตกต่างกัน

งานวิจัยที่เป็นกลุ่มตัวอย่างในกลุ่มที่มีความเสี่ยงเกิดความคลาดเคลื่อนในงานวิจัย สามารถแจกแจงการละเมิดข้อตกลงในแต่ละกลุ่มสถิติได้ดังรายละเอียดในตารางที่ 5.23

ตารางที่ 5.23 แสดงจำนวนและร้อยละของกลุ่มตัวอย่างที่มีการละเมิดข้อตกลงระหว่างกลุ่มสถิติจำแนกตามตัวแปรหุ่น

ตัวแปรหุ่น	สถิติ		รวม
	แบบมีพารามิเตอร์	แบบไม่มีพารามิเตอร์	
เครื่องมือวัด	21.49 (26)	26.45 (32)	47.93 (58)
ระดับการวัด	26.45 (32)	28.93 (25)	55.37 (67)
การเลือกตัวอย่าง	30.58 (37)	11.57 (14)	42.15 (51)
ขนาดตัวอย่าง	36.36 (44)	31.40 (38)	67.77 (82)
ประชากร	4.96 (6)	3.31 (4)	8.26 (10)
การตรวจสอบข้อมูล	64.46 (78)	33.88 (41)	98.35 (119)
รวม	66.12 (80)	33.88 (41)	100.00 (121)

จากตารางที่ 5.23 กลุ่มสถิติแบบมีพารามิเตอร์ กำหนดให้การละเมิดข้อตกลงทุกตัวแปรที่มีโอกาสทำให้มีความเสี่ยงเกิดความคลาดเคลื่อนในงานวิจัย โดยตัวแปรที่มีการละเมิดข้อตกลงมากที่สุด คือ ไม่มีการตรวจสอบข้อมูล รองลงมา คือ การกำหนดขนาดตัวอย่างโดยไม่ใช้หลักสถิติ การเลือกตัวอย่างแบบไม่ใช้หลักความน่าจะเป็น มีระดับการวัดไม่ใช่ค่าต่อเนื่อง ไม่มีการตรวจสอบเครื่องมือวัด และน้อยที่สุด คือ ที่มาของประชากรเป็นแบบไม่ทราบจำนวน ส่วนในกลุ่มสถิติแบบไม่มีพารามิเตอร์ เนื่องจากเป็นสถิติที่เป็นอิสระจากข้อตกลง (assumption-free tests) จึงกำหนดเงื่อนไขเดียวที่มีโอกาสทำให้มีความเสี่ยงเกิดความคลาดเคลื่อนในงานวิจัย คือ ไม่มีการตรวจสอบเครื่องมือวัด

จากผลการวิเคราะห์ความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยในเบื้องต้น โดยเริ่มจากการพิจารณาจากตัวแปรที่น่ามีผลทำให้เกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย พบว่า ตัวแปรที่น่าจะมีผลทำให้เกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากที่สุด คือ การตรวจสอบข้อมูล รองลงมา คือ สมมติฐาน ขนาดตัวอย่าง ระดับการวัด สถิติ การเลือกตัวอย่าง เครื่องมือวัด และน้อยที่สุด คือ ประชากร การวิเคราะห์แบ่งกลุ่ม พบว่า ตัวอย่างที่เป็นงานวิจัยมีรูปแบบการวิธีการดำเนินการวิจัยที่คล้ายคลึงกันจำนวน 30 กลุ่ม มีกลุ่มตัวอย่างงานวิจัยที่มีการทดสอบสมมติฐานและมีโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย 26 กลุ่ม จำนวน 225 เรื่อง เมื่อนำมาวิเคราะห์และกำหนดให้เป็นงานวิจัยที่มีโอกาสที่จะมีความเสี่ยงและไม่มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยและทดสอบสัดส่วนเชิงสถิติพบว่า มีจำนวนไม่แตกต่างกัน

การวิเคราะห์การทำนายความเสี่ยงการเกิดความคลาดเคลื่อนในงานวิจัย

จากการวิเคราะห์ความเสี่ยงการเกิดความคลาดเคลื่อนของสมมติฐานที่เป็นผลมาจากการละเมิดข้อตกลงเบื้องต้นที่ใช้ในวิเคราะห์ข้อมูลด้วยสถิติแต่ละประเภท โดยใช้ตัวแปรหุ่นตามกรอบแนวคิดของการวิจัย คือ เครื่องมือวัด ระดับการวัด ประชากร การกำหนดขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล พบว่า มีความเสี่ยงและไม่มีความเสี่ยงในสัดส่วนที่ไม่แตกต่างกัน จึงนำตัวแปรหุ่นชุดดังกล่าวมาวิเคราะห์การถดถอยเชิงโลจิสติกพหุ (Multiple Logistic Regression Analysis: MLRA) เพื่อนำไปวิเคราะห์การพยากรณ์หรือทำนาย (predicted) ความน่าจะเป็น (probability) ของโอกาสที่จะเกิด (odds) ความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย (risk) จากอัตราส่วนระหว่างมีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยกับไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัย โดยมีสมการดังนี้

$$\frac{1}{1+e^{-(B_0+B_1X_1\ldots+B_nX_n)}}$$

การวิเคราะห์ความเหมาะสมของแบบจำลอง

การวิเคราะห์ความเหมาะสมของแบบจำลองในการทำนายโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยจากตัวแปรทำนาย 6 ตัวแปร คือ เครื่องมือวัด ระดับการวัด ประชากร ขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล เพื่อให้ได้ค่าสัมประสิทธิ์การถดถอยของสมการถดถอยเชิงโลจิสติกสำหรับใช้คำนวณความน่าจะเป็นสูงสุดในการทำนายโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย โดยมีแบบจำลองและวิธีการวิเคราะห์ดังต่อไปนี้

1. แบบจำลอง 1 เป็นแบบจำลองที่ใช้การวิเคราะห์แบบนำตัวแปรทำนาย 6 ตัวแปร นำเข้าวิเคราะห์ในแบบจำลองพร้อมกันทั้งหมด (Enter Method) เพื่อทดสอบความเหมาะสมของตัวแบบในการทำนายและผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย

2. แบบจำลอง 2 เป็นแบบจำลองที่ใช้วิธีการวิเคราะห์แบบคัดเลือกตัวแปรทำนายออกจากแบบจำลองอย่างเป็นขั้นตอน (Backward Stepwise: Likelihood Ratio) โดยคัดเลือกตัวแปรทำนายที่ไม่มีผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยออกจากแบบจำลองทีละ 1 ตัวแปร จนได้แบบจำลองที่เหมาะสมในการทำนาย

3. แบบจำลอง 3 เป็นแบบจำลองที่ใช้วิธีการเคราะห์แบบคัดเลือกตัวแปรทำนายเข้าแบบจำลองอย่างเป็นขั้นตอน (Forward Stepwise: Likelihood Ratio) โดยคัดเลือกตัวแปรทำนายที่มีผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมากที่สุดอย่างมีนัยสำคัญทางสถิติเข้าแบบจำลองทีละ 1 ตัวแปร จนได้แบบจำลองในการทำนายได้ถูกต้องมากที่สุด

การวิเคราะห์เพื่อคัดเลือกตัวแปรทำนายสำหรับสร้างสมการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย มีรายละเอียดดังต่อไปนี้

1. การแปลงค่าตัวแปรหุ่น (Recoding Dummy Variables)

ก่อนการวิเคราะห์ได้ทำการแปลงค่าตัวแปรหุ่นที่เป็นตัวแปรทำนาย โดยให้ค่า 0 = ไม่ละเมิดข้อตกลง และ 1 = ละเมิดข้อตกลง โดยมีรายละเอียดดังตารางที่ 5.24

ตารางที่ 5.24 แสดงการแปลงค่าตัวแปรทำนาย

ตัวแปรทำนาย	ค่าตัวแปรและความหมาย	ค่าตัวแปร
เครื่องมือวัด (X_1)	0 = ไม่มีการทดสอบ	1
	1 = มีการทดสอบ	0
ระดับการวัด (X_2)	0 = ค่าไม่ต่อเนื่อง	1
	1 = ค่าต่อเนื่อง	0
ประชากร (X_3)	0 = ไม่ทราบจำนวน	1
	1 = ทราบจำนวน	0
ขนาดตัวอย่าง (X_4)	0 = ไม่คำนวนด้วยวิธีการทางสถิติ	1
	1 = คำนวนด้วยวิธีการทางสถิติ	0
การเลือกตัวอย่าง (X_5)	0 = ไม่ใช้หลักความน่าจะเป็น	1
	1 = ใช้หลักความน่าจะเป็น	0
การตรวจสอบข้อมูล (X_6)	0 = ไม่มีการตรวจสอบ	1
	1 = มีการตรวจสอบ	0

2. การทดสอบภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรทำนาย (Multicollinearity Analysis)

เนื่องจากข้อมูลตัวแปรทำนาย 6 ตัว คือ เครื่องมือวัด ระดับการวัด การเลือกตัวอย่าง ขนาดตัวอย่าง ประชากร และการตรวจสอบข้อมูล มีค่าเป็น 0 กับ 1 และมีข้อมูลจำนวน 225 ชุด (case) สอดคล้องกับข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยเชิงโลจิสติกพหุ แต่ยังไม่ได้ทำการทดสอบภาวะร่วมเชิงเส้นตรง

ระหว่างตัวแปรทำนาย จึงทำการตรวจสอบความสัมพันธ์เบื้องต้นระหว่างตัวแปรด้วยการวิเคราะห์สัมประสิทธิ์สหสัมพันธ์ (correlation coefficient) แบบเพียร์สัน (Pearson) โดยมีรายละเอียดดังในตารางที่ 5.25

ตารางที่ 5.25 แสดงค่าการทดสอบความสัมพันธ์ระหว่างตัวแปรทำนาย

ตัวแปรทำนาย	เครื่องมือวัด	ระดับการวัด	ประชากร	ขนาดตัวอย่าง	การเลือกตัวอย่าง	การตรวจสอบข้อมูล
เครื่องมือวัด	1.00	0.187**	0.152*	0.361**	0.007	0.059
ระดับการวัด		1.00	0.048	0.153*	-0.086	0.102
ประชากร			1.00	0.238**	0130*	0.027
ขนาดตัวอย่าง				1.00	0.353**	-0.084
การเลือกตัวอย่าง					1.00	-0.110
การตรวจสอบข้อมูล						1.00

* มีนัยสำคัญทางสถิติที่ระดับ 0.05 (แบบสองหาง) ** มีนัยสำคัญทางสถิติที่ระดับ 0.01 (แบบสองหาง)

จากตารางที่ 5.25 แสดงให้เห็นว่า ตัวแปรทำนายมีความสัมพันธ์เชิงบวกอย่างมีนัยสำคัญทางสถิติ 7 คู่ คือ เครื่องมือวัดกับระดับการวัด เครื่องมือวัดกับประชากร เครื่องมือวัดกับขนาดตัวอย่าง ระดับการวัดกับขนาดตัวอย่าง ประชากรกับขนาดตัวอย่าง ประชากรกับการเลือกตัวอย่าง และขนาดตัวอย่างกับการเลือกตัวอย่าง

เมื่อพิจารณาจากค่าสัมประสิทธิ์สหสัมพันธ์พบว่า ตัวแปรทำนายทั้งหมดไม่มีตัวแปรคู่ใดมีค่าสัมประสิทธิ์สหสัมพันธ์มากกว่า 0.7 จึงไม่น่าเกิดภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรทำนาย และจากการทดสอบภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรทำนายได้ค่าสถิติภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรทำนาย คือ ค่าความคลาดเคลื่อนที่ยอมรับได้ (Tolerance) และค่าองค์ประกอบการกระจายของความแปรปรวน (Variance Inflation Factor: VIF) โดยมีรายละเอียดดังในตารางที่ 5.26

ตารางที่ 5.26 แสดงค่าการทดสอบภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรทำนาย

ตัวแปรทำนาย	ค่า Tolerance	ค่า VIF
เครื่องมือวัด	0.848	1.180
ระดับการวัด	0.940	1.064
ประชากร	0.932	1.073
ขนาดตัวอย่าง	0.702	1.425
การเลือกตัวอย่าง	0.833	1.200
การตรวจสอบข้อมูล	0.965	1.037

จากตารางที่ 5.26 ค่าความคลาดเคลื่อนที่ยอมรับได้ (Tolerance) มีค่าเข้าใกล้ 1 แสดงว่าตัวแปรเป็นอิสระจากกัน และค่าองค์ประกอบการกระจายของความแปรปรวน (Variance Inflation Factor: VIF) มีค่าไม่ถึง 10.00 แสดงว่าตัวแปรทำนายเป็นอิสระจากกัน ไม่เกิดภาวะร่วมเชิงเส้นตรงระหว่างตัวแปรทำนาย

3. การวิเคราะห์ความสอดคล้องโดยรวมของแบบจำลอง (Model Summary Analysis)

ผลการทดสอบความเหมาะสมของตัวแปรทำนายทั้งหมดที่ใช้ในแบบจำลองหรือการทดสอบแบบออมนิบัส (omnibus test) เพื่อค้นหากลุ่มตัวแปรทำนายที่สามารถอธิบาย (explanation) หรือมีอิทธิพลต่อ (effects) ความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย มีรายละเอียดดังในตารางที่ 5.27

ตารางที่ 5.27 แสดงค่าการทดสอบความเหมาะสมของตัวแปรทำนาย

แบบจำลอง		Chi-square	df	Sig.
1	Step	74.578	6	0.000
	Block	74.578	6	0.000
	Model	74.578	6	0.000
2	Step	-2.532	1	0.112
	Block	69.549	3	0.000
	Model	69.549	3	0.000
3	Step	7.258	1	0.007
	Block	66.447	2	0.000
	Model	66.447	2	0.000

จากตารางที่ 5.27 การทดสอบแบบจำลอง 1 แบบจำลอง 2 และแบบจำลอง 3 ด้วยค่าไคสแควร์ (Chi-square) มีนัยสำคัญทางสถิติ (Sig. < 0.05) ทุกแบบจำลอง

เมื่อพิจารณาโดยรวม แสดงว่า แต่ละแบบจำลองมีตัวแปรทำนายในแบบจำลองอย่างน้อย 1 ตัวมีความแตกต่างจากค่า 0 หรือพบว่า แต่ละแบบจำลองมีตัวแปรทำนายบางตัวแปรที่มีผลกระทบต่อการเปลี่ยนแปลงความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย

4. การวิเคราะห์ความสามารถในการทำนายของแบบจำลอง (Model Prediction Analysis)

ผลการวิเคราะห์ค่าสัมประสิทธิ์การกำหนดหรือการตัดสินใจ (coefficient of determination) ของแบบจำลอง เพื่อเปรียบเทียบความสามารถของตัวแปรทำนาย (ตัวแปรอิสระ) ในการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย (ตัวแปรตามหรือตัวแปรถูกทำนาย) ด้วยค่า R^2 เทียม (pseudo R^2) โดยวิธีของคอกซ์และสเนลล์ (Cox and Snell R^2: RCS^2) และของนาเกลเกอร์กี้ (Nagelkerke R^2: RN^2) มีรายละเอียดดังในตารางที่ 5.28

ตารางที่ 5.28 แสดงค่าการทดสอบความสามารถในการทำนายของแบบจำลอง

แบบจำลอง	ค่า -2 log	ค่า RCS^2	ค่า RN^2
1	236.053	0.282	0.377
2	241.082	0.266	0.355
3	244.184	0.256	0.342

จากตารางที่ 5.28 ผลการวิเคราะห์ความสามารถของตัวแปรทำนายในการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยได้ค่าคอกซ์และสเนลล์ (RCS^2) และนาเกลเกอร์กี้ (RN^2) มีรายละเอียดดังต่อไปนี้

แบบจำลอง 1 มีค่าสัมประสิทธิ์การตัดสินใจของแบบจำลองโดยวิธีของคอกซ์และสเนลล์ประมาณร้อยละ 28.2 และของนาเกลเกอร์กี้ประมาณร้อยละ 37.7

แบบจำลอง 2 มีค่าสัมประสิทธิ์การตัดสินใจของแบบจำลองโดยวิธีของคอกซ์และสเนลล์ประมาณร้อยละ 26.6 และของนาเกลเกอร์กี้ประมาณร้อยละ 35.5

แบบจำลอง 3 มีค่าสัมประสิทธิ์การตัดสินใจของแบบจำลองโดยวิธีของคอกซ์และสเนลล์ประมาณร้อยละ 25.6 และของนาเกลเกอร์กี้ประมาณร้อยละ 34.2

เมื่อเปรียบเทียบจากค่าสัมค่าสัมประสิทธิ์ของแบบจำลองทั้งของคอกซ์และสเนลล์ และของนาเกลเกอร์กี้ แสดงว่า แบบจำลอง 1 ตัวแปรทำนายมีความสามารถในการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยดีกว่าแบบจำลอง 2 และแบบจำลอง 3 และแบบจำลอง 2 ตัวแปรทำนายมีความสามารถในการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยดีกว่าแบบจำลอง 3 โดยแบบจำลอง 1 ตัวแปรทำนายมีความสามารถในการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมากที่สุด และแบบจำลอง 3 ตัวแปรทำนายมีความสามารถในการทำนายน้อยที่สุด

เนื่องจากค่าสัมประสิทธิ์ของแบบจำลองไม่สามารถบอกได้ว่า แบบจำลองมีความเหมาะสมดีกับข้อมูลหรือไม่ จึงทำการวิเคราะห์เพื่อทดสอบความสอดคล้องของแบบจำลองกับข้อมูลในขั้นต่อไป

5. การวิเคราะห์ความสอดคล้องของแบบจำลองกับข้อมูล (Goodness of Fit)

ผลการทดสอบความแนบสนิทหรือความกลมกลืนของแบบจำลองด้วยวิธีของฮอสเมอร์และเลมส์โชว์ (Hosmer and Lemeshow) เพื่อทดสอบความสัมพันธ์หรือความสอดคล้องกันระหว่างข้อมูลหรือค่าจริงที่ได้จากการสังเกต (observed value) จากกลุ่มตัวอย่าง กับข้อมูลหรือค่าคาดหมาย (expected value) จากความน่าจะเป็นเชิงทฤษฎีของแบบจำลอง มีรายละเอียดดังในตารางที่ 5.29

ตารางที่ 5.29 แสดงค่าการทดสอบความสอดคล้องของแบบจำลองกับข้อมูล

แบบจำลอง	Chi-square	df	Sig.
1	23.605	6	0.001
2	1.159	3	0.763
3	1.980	2	0.372

จากตารางที่ 5.29 ผลการทดสอบความสอดคล้องของแบบจำลองกับข้อมูลด้วยวิธีของฮอสเมอร์และเลมส์โชว์ มีรายละเอียดดังต่อไปนี้

แบบจำลอง 1 ที่ประกอบด้วยตัวแปรทำนาย 6 ตัว ได้ผลการทดสอบความสอดคล้องของข้อมูลมีนัยสำคัญทางสถิติ (Sig. < 0.05) แสดงว่า เป็นแบบจำลองที่มีค่าสังเกตแตกต่างกับค่าคาดหมาย (ข้อมูลไม่มีความสอดคล้องกัน)

แบบจำลอง 2 ที่ประกอบด้วยตัวแปรทำนาย 3 ตัวแปร ได้ผลการทดสอบความสอดคล้องของข้อมูลไม่มีนัยสำคัญทางสถิติ (Sig. >= 0.05) แสดงว่า เป็นแบบจำลองที่มีค่าสังเกตไม่แตกต่างกับค่าคาดหมาย (ข้อมูลมีความสอดคล้องกัน)

แบบจำลอง 3 ที่ประกอบด้วยตัวแปรทำนาย 2 ตัวแปร ได้ผลการทดสอบความสอดคล้องของข้อมูลไม่มีนัยสำคัญทางสถิติ (Sig. >= 0.05) แสดงว่า เป็นแบบจำลองที่มีค่าสังเกตไม่แตกต่างกับค่าคาดหมาย (ข้อมูลมีความสอดคล้องกัน)

6. การวิเคราะห์ความถูกต้องของตัวแปรถูกทำนาย

ผลการวิเคราะห์ความถูกต้องของการแบ่งกลุ่ม (correctly classified) ความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย (predicted variables) หรือประสิทธิภาพของแบบจำลองในการทำนายความถูกต้องของตัวแปรถูกทำนาย (ตัวแปรตามหรือตัวแปรพยากรณ์) มีรายละเอียดดังในตารางที่ 5.30

ตารางที่ 5.30 แสดงค่าการทดสอบของแบบจำลองในการทำนายความถูกต้องของตัวแปรถูกทำนาย

แบบจำลอง	ความเสี่ยง	การทำนาย		
		ความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย		ความถูกต้อง (ร้อยละ)
		ไม่มี (No)	มี (Yes)	
1	ไม่มี (No)	83	21	79.8
	มี (Yes)	37	84	69.4
	รวม (ร้อยละ)			74.2
2	ไม่มี (No)	71	33	68.3
	มี (Yes)	31	90	74.4
	รวม (ร้อยละ)			71.6
3	ไม่มี (No)	64	40	61.5
	มี (Yes)	30	91	75.2
	รวม (ร้อยละ)			68.9

ค่าของจุดแบ่ง (cut value) = 0.5

จากตารางที่ 5.30 แบบจำลองมีประสิทธิภาพในการทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยได้ดังต่อไปนี้

แบบจำลอง 1 สามารถทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยโดยรวมได้ถูกต้องประมาณร้อยละ 74.2 ถ้ามีการละเมิดข้อตกลง (ค่าตัวแปรทำนาย = 1) และคำนวณค่าความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อในงานวิจัยได้มากกว่า 0.5 (Prob. > 0.5) สามารถทำนายได้ถูกต้องประมาณร้อยละ 69.4 แต่ถ้าไม่มีการละเมิดข้อตกลง (ค่าตัวแปรทำนาย = 0) และคำนวณความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อในงานวิจัยได้น้อยกว่าหรือเท่ากับ 0.5 (Prob. <= 0.5) สามารถทำนายได้ถูกต้องประมาณร้อยละ 79.8

แบบจำลองที่ 2 สามารถทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยโดยรวมได้ถูกต้องประมาณร้อยละ 71.6 ถ้ามีการละเมิดข้อตกลง (ค่าตัวแปรทำนาย = 1) และ

คำนวณค่าความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อในงานวิจัยได้มากกว่า 0.5 (Prob. > 0.5) สามารถทำนายได้ถูกต้องประมาณร้อยละ 74.4 แต่ถ้าไม่มีการละเมิดข้อตกลง (ค่าตัวแปรทำนาย = 0) และคำนวณความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อในงานวิจัยได้น้อยกว่าหรือเท่ากับ 0.5 (Prob. <= 0.5) สามารถทำนายได้ถูกต้องประมาณร้อยละ 68.3

แบบจำลองที่ 3 สามารถทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยโดยรวมได้ถูกต้องประมาณร้อยละ 68.9 ถ้ามีการละเมิดข้อตกลง (ค่าตัวแปรทำนาย = 1) และคำนวณค่าความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อในงานวิจัยได้มากกว่า 0.5 (Prob. > 0.5) สามารถทำนายได้ถูกต้องประมาณร้อยละ 75.2 แต่ถ้าไม่มีการละเมิดข้อตกลง (ค่าตัวแปรทำนาย = 0) และคำนวณความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อในงานวิจัยได้น้อยกว่าหรือเท่ากับ 0.5 (Prob. <= 0.5) สามารถทำนายได้ถูกต้องประมาณร้อยละ 61.5

7. การวิเคราะห์คัดเลือกแบบจำลอง

ผลการวิเคราะห์การถดถอยเชิงโลจิสติกพหุเพื่อคัดเลือกตัวแปรทำนายสำหรับสร้างสมการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยมี 3 วิธี คือ วิธีการนำตัวแปรเข้าแบบจำลองพร้อมกันทั้งหมด (Enter Method) หรือ แบบจำลอง 1 วิธีการคัดเลือกตัวแปรทำนายออกจากแบบจำลองอย่างเป็นขั้นตอนโดยใช้ค่าสัดส่วนความควรจะเป็น (Backward Stepwise: Likelihood Ratio) หรือแบบจำลอง 2 และวิธีการคัดเลือกตัวแปรทำนายเพิ่มเข้าแบบจำลองอย่างเป็นขั้นตอนโดยใช้ค่าสัดส่วนความควรจะเป็น (Forward Stepwise: Likelihood Ratio) หรือ แบบจำลอง 3

ผลการวิเคราะห์ของแบบจำลอง 1 สมการที่ประกอบด้วยตัวแปรทำนายทั้ง 6 ตัวแปร คือ เครื่องมือวัด ระดับการวัด ประชากร ขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล ไม่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ มีรายละเอียดดังในตารางที่ 5.31

ตารางที่ 5.31 แสดงค่าตัวแปรทำนายความเสี่ยงในการเกิดความคลาดเคลื่อนในงานวิจัยของแบบจำลอง 1

ตัวแปรทำนาย	ค่าตัวแปรทำนายในสมการ						
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for Exp(B)
เครื่องมือวัด (X_1)	2.811	0.522	28.987	1	0.000*	16.631	5.997-46.278
ระดับการวัด (X_2)	-0.338	0.323	1.092	1	0.296	0.713	0.379-1.344
ประชากร (X_3)	-1.077	0.664	2.630	1	0.105	0.341	0.93-1.252
ขนาดตัวอย่าง (X_4)	1.169	0.378	9.568	1	0.002*	3.219	1.535-6.753
การเลือกตัวอย่าง (X_5)	-0.458	0.370	1.537	1	0.215	0.632	0.306-1.305
การตรวจสอบข้อมูล (X_6)	-21.079	28420.696	0.000	1	0.999	0.000	0.00-∞
ค่าคงที่ (Constant)	20.492	28420.696	0.000	1	0.999	793573453.9	

* มีนัยสำคัญทางสถิติน้อยกว่า 0.05

จากตารางที่ 5.31 แม้ว่าตัวแปรทำนายทั้ง 6 ตัวแปร ไม่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ แต่มีตัวแปรทำนาย 2 ตัวแปรที่มีผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ คือ เครื่องมือวัด และขนาดตัวอย่าง

ผลการวิเคราะห์ของแบบจำลอง 2 สมการที่ได้จากการคัดเลือกตัวแปรทำนายออกจากตัวแปรทำนาย 6 ตัวแปร จนได้มีตัวแปรทำนายที่เหมาะสมเหลืออยู่ในสมการ 3 ตัวแปร และมีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ รายละเอียดดังในตารางที่ 5.32

ตารางที่ 5.32 แสดงค่าตัวแปรทำนายความเสี่ยงในการเกิดความคลาดเคลื่อนในงานวิจัยของแบบจำลอง 2

ตัวแปรทำนาย	ค่าตัวแปรทำนายในสมการ						
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for Exp(B)
เครื่องมือวัด (X_1)	2.754	0.518	28.235	1	0.000*	15.703	5.686-43.362
ประชากร (X_3)	-1.134	0.656	2.989	1	0.084	0.322	0.89-1.164
ขนาดตัวอย่าง (X_4)	0.976	0.326	8.956	1	0.003*	2.653	1.400-5.026
ค่าคงที่ (Constant)	-0.822	0.220	13.999	1	0.000	0.440	

* มีนัยสำคัญทางสถิติน้อยกว่า 0.05

จากตารางที่ 5.32 ตัวแปรทำนาย 3 ตัวแปร ที่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ คือ เครื่องมือวัด ประชากร และขนาดตัวอย่าง

งานวิจัยที่มีการละเมิดข้อตกลงด้านเครื่องมือวัด (ไม่มีการตรวจสอบเครื่องมือวัด) เมื่อเทียบกับไม่ละเมิดข้อตกลงด้านเครื่องมือวัด (มีการตรวจสอบเครื่องมือวัด) มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่าประมาณ 16 เท่า (ค่า B = 2.754 และ Exp(B) หรือ Odds Ratio: OR = 15.703)

งานวิจัยที่มีการละเมิดข้อตกลงด้านประชากร (ไม่ทราบจำนวนประชากร) เมื่อเทียบกับไม่ละเมิดข้อตกลงด้านประชากร (ทราบจำนวนประชากร) มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย น้อยกว่าประมาณ 0.322 เท่า (ค่า B = -1.134 และ ค่า Exp(B) หรือ Odds Ratio: OR = 0.322)

งานวิจัยที่มีการละเมิดข้อตกลงด้านขนาดตัวอย่าง (ไม่ใช้วิธีการทางสถิติ) เมื่อเทียบกับไม่ละเมิดข้อตกลงด้านขนาดตัวอย่าง (ใช้วิธีการทางสถิติ) มีความสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่าประมาณ 2.653 เท่า (ค่า B = 0.976 และค่า Exp(B) หรือ Odds Ratio: OR = 2.653)

ผลการวิเคราะห์ของแบบจำลอง 3 สมการที่ได้จากคัดเลือกตัวแปรทำนายที่สามารถทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติเข้าสมการครั้งละ 1 ตัวแปร สามารถคัดเลือกตัวแปรทำนายเข้าสมการได้ 2 ตัวแปร รายละเอียดดังในตารางที่ 5.33

ตารางที่ 5.33 แสดงค่าตัวแปรทำนายความเสี่ยงในการเกิดความคลาดเคลื่อนในงานวิจัยของแบบจำลอง 3

ตัวแปรทำนาย	ค่าตัวแปรทำนายในสมการ						
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for Exp(B)
เครื่องมือวัด (X_1)	2.643	0.502	27.749	1	0.000*	14.059	5.258-37.589
ขนาดตัวอย่าง (X_4)	0.843	0.314	7.212	1	0.007*	2.322	1.256-4.296
ค่าคงที่ (Constant)	-0.823	0.219	14.142	1	0.000*	0.439	

* มีนัยสำคัญทางสถิติน้อยกว่า 0.05

จากตารางที่ 5.33 ตัวแปรทำนาย 2 ตัวแปร ที่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ คือ เครื่องมือวัด และขนาดตัวอย่าง

งานวิจัยที่มีการละเมิดข้อตกลงด้านเครื่องมือวัด (ไม่มีการตรวจสอบเครื่องมือวัด) เมื่อเทียบกับไม่ละเมิดข้อตกลงด้านเครื่องมือวัด (มีการตรวจสอบเครื่องมือวัด) มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่าประมาณ 14 เท่า (ค่า B = 2.643 และค่า Exp(B) หรือ Odds Ratio: OR = 14.059)

งานวิจัยที่มีการละเมิดข้อตกลงด้านขนาดตัวอย่าง (ไม่ใช้วิธีการทางสถิติ) เมื่อเทียบกับไม่ละเมิดข้อตกลงด้านขนาดตัวอย่าง (ใช้วิธีการทางสถิติ) มีความสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่าประมาณ 2 เท่า (ค่า B = 0.843 และค่า Exp(B) หรือ Odds Ratio: OR = 2.322)

จากผลการการวิเคราะห์แบบจำลองทั้ง 3 แบบจำลอง มีตัวแปรทำนายที่มีผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติเหมือนกัน คือ เครื่องมือวัด และขนาดตัวอย่าง แต่แบบจำลองทั้ง 3 แบบจำลอง มี 1 แบบจำลองที่ตัวแปรทำนายไม่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ คือ แบบจำลองที่ 1 และมี 2 แบบจำลองที่มีตัวแปรทำนายมีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ คือ แบบจำลอง 2 และแบบจำลอง 3

8. การทำนายความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย

จากการวิเคราะห์คัดเลือกแบบจำลอง มีแบบจำลองที่มีตัวแปรทำนายมีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ คือ แบบจำลอง 2 และแบบจำลอง 3 จึงนำมาคำนวณเปรียบเทียบค่าความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย โดยมีรายละเอียดดังต่อไปนี้

1. แบบจำลอง 2 สมการมีตัวแปรทำนาย 3 ตัวแปร คือ เครื่องมือวัด (X_1) ประชากร (X_3) และขนาดตัวอย่าง (X_4) สามารถคำนวณค่าความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยได้จากสมการดังนี้

$$P(R) = \frac{1}{1+e^{-(B_0+B_1X_1+B_2X_3+B_3X_4)}}$$

เมื่อกำหนดว่า มีการละเมิดข้อตกลงด้านเครื่องมือวัด (ไม่มีการตรวจสอบเครื่องมือวัด -> $X_1 = 1$) มีการละเมิดข้อตกลงด้านประชากร (ไม่ทราบจำนวนประชากร -> $X_3 = 1$) และมีการละเมิดข้อตกลงด้านขนาดตัวอย่าง (ไม่ใช้วิธีการทางสถิติ -> $X_4 = 1$) คำนวณความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยได้ดังนี้

$$P(R) = \frac{1}{1+2.71828^{-(-0.822+2.754(1)-1.134(1)+0.843(1))}} = 0.85$$

จากการคำนวณความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยจากตัวแปรทำนายที่มีการละเมิดข้อตกลงด้านเครื่องมือวัด ด้านประชากร และด้านขนาดตัวอย่าง ได้ค่ามากกว่า 0.5 (Prob. = 0.85) จึงทำนายได้ว่า มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย

เมื่อกำหนดว่า ไม่มีการละเมิดข้อตกลงการตรวจสอบด้านเครื่องมือวัด (มีการตรวจสอบเครื่องมือวัด -> $X_1 = 0$) ไม่มีการละเมิดข้อตกลงด้านประชากร (ทราบจำนวนประชากร -> $X_3 = 0$) และไม่มีการละเมิดข้อตกลงด้านขนาดตัวอย่าง (ใช้วิธีการทางสถิติ -> $X_4 = 0$) คำนวณความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยได้ดังนี้

$$P(R) = \frac{1}{1+2.71828^{-(-0.822+2.754(0)-1.134(0)+0.843(0))}} = 0.31$$

จากการคำนวณความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยจากตัวแปรทำนายที่ไม่มีการละเมิดข้อตกลงด้านเครื่องมือวัด ด้านประชากร และด้านขนาดตัวอย่าง ได้ค่าน้อยกว่า 0.5 (Prob. = 0.31) จึงทำนายได้ว่า ไม่มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย

2. แบบจำลอง 3 สมการมีตัวแปรทำนาย 2 ตัวแปร คือ เครื่องมือวัด (X_1) และขนาดตัวอย่าง (X_4) สามารถคำนวณค่าความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยได้จากสมการดังนี้

$$P(R) = \frac{1}{1+e^{-(B_0+B_1X_1+B_2X_4)}}$$

เมื่อกำหนดว่า มีการละเมิดข้อตกลงด้านเครื่องมือวัด (ไม่มีการตรวจสอบเครื่องมือวัด -> $X_1 = 1$ และมีการละเมิดข้อตกลงด้านขนาดตัวอย่าง (ไม่ใช้วิธีการทางสถิติ -> $X_4 = 1$) คำนวณความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยได้ดังนี้

$$P(R) = \frac{1}{1+2.71828^{-(-0.823+2.643(1)+0.843(1))}} = 0.93$$

จากการคำนวณความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยจากตัวแปรทำนายที่มีการละเมิดข้อตกลงด้านเครื่องมือวัด และด้านขนาดตัวอย่าง ได้ค่ามากกว่า 0.5 (Prob. = 0.93) จึงทำนายได้ว่า มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย

เมื่อกำหนดว่า ไม่มีการละเมิดข้อตกลงด้านเครื่องมือวัด (มีการตรวจสอบเครื่องมือวัด -> $X_1 = 0$ และไม่มีการละเมิดข้อตกลงด้านขนาดตัวอย่าง (ใช้วิธีการทางสถิติ -> $X_4 = 0$) คำนวณความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยได้ดังนี้

$$P(R) = \frac{1}{1+2.71828^{-(-0.823+2.643(0)+0.843(0))}} = 0.31$$

จากการคำนวณความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยจากตัวแปรทำนายที่มีการละเมิดข้อตกลงด้านเครื่องมือวัด และด้านขนาดตัวอย่าง ได้ค่าน้อยกว่า 0.5 (Prob. = 0.31) จึงทำนายได้ว่า ไม่มีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย

จากการเปรียบเทียบความน่าจะเป็นของโอกาสที่จะเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยระหว่างแบบจำลอง 2 กับแบบจำลอง 3 แสดงให้เห็นว่า แบบจำลอง 3 มีค่าความน่าจะเป็นของโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่าแบบจำลอง 2

แบบจำลอง 3 มีตัวแปรทำนายที่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ 2 ตัวแปร คือ เครื่องมือวัด และขนาดตัวอย่าง

ดังนั้นเพื่อให้สามารถอธิบายได้ว่า เมื่อตัวแปรทำนายตัวแปรใดตัวแปรหนึ่งมีค่าเพิ่มขึ้น ขณะที่ตัวแปรทำนายอื่นๆ คงที่ จะทำให้ความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยเปลี่ยนแปลงไปอย่างไร จึงทำการวิเคราะห์หาผลกระทบส่วนเพิ่ม (Marginal Effects: ME) โดยมีรายละเอียดดังในตารางที่ 5.34

ตารางที่ 5.34 แสดงค่าประมาณการผลกระทบส่วนเพิ่มของตัวแปรทำนาย

ตัวแปรทำนาย	Marginal Effects (dy/dx)	S. E.	Z	Sig.
เครื่องมือวัด (X_1)	0.4993	0.0570	8.747	0.000*
ขนาดตัวอย่าง (X_4)	0.2020	0.0736	2.742	0.006*

* มีนัยสำคัญทางสถิติน้อยกว่า 0.05

จากตารางที่ 5.34 ผลการวิเคราะห์ค่าประมาณการผลกระทบส่วนเพิ่มของตัวแปรทำนาย มีรายละเอียดดังนี้

1. ตัวแปรทำนายด้านเครื่องมือวัด อธิบายได้ว่า เมื่อกำหนดให้ปัจจัยอื่นๆคงที่ งานวิจัยที่ละเมิดข้อตกลงโดยไม่มีการทดสอบเครื่องมือวัด (เครื่องมือวัด = 1) มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการ

วิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงโดยมีการทดสอบเครื่องมือวัด (เครื่องมือวัด = 0) ประมาณร้อยละ 49.93 (ME = 0.4993)

2. ตัวแปรทำนายด้านขนาดตัวอย่าง อธิบายได้ว่า เมื่อกำหนดให้ปัจจัยอื่นๆคงที่ งานวิจัยที่ละเมิดข้อตกลงโดยไม่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง (ขนาดตัวอย่าง = 1) มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงโดยใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง (ขนาดตัวอย่าง = 0) ประมาณร้อยละ 20.20 (ME = 0.2020)

จากผลการวิเคราะห์สรุปได้ว่า แบบจำลองที่ประกอบด้วยตัวแปรทำนาย 2 ตัวแปร คือ เครื่องมือวัดและการคำนวณขนาดตัวอย่าง สามารถทำนายความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยได้ประมาณร้อยละ 34.2 (Nagelkerke R^2 = 0.342) และทำนายความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยได้ถูกต้องประมาณร้อยละ 75.2 โดยสามารถทำนายได้ว่า ตัวอย่างงานวิจัยที่มีการละเมิดข้อตกลงด้านเครื่องมือวัด และการคำนวณขนาดตัวอย่าง มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัย (Prob. = 0.93) โดยตัวแปรทำนายด้านเครื่องมือวัดมีผลกระทบต่อความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยมากกว่าตัวแปรด้านการคำนวณขนาดตัวอย่าง กล่าวคือ งานวิจัยที่มีการละเมิดข้อตกลงด้านเครื่องมือวัดมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงด้านเครื่องมือวัดประมาณ 14 เท่า (OR = 14.059) หรือประมาณร้อยละ 49.93 (ME = 0.4993) และงานวิจัยที่มีการละเมิดข้อตกลงด้านขนาดตัวอย่างมีความสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงด้านขนาดตัวอย่างประมาณ 2 เท่า (OR = 2.322) หรือประมาณร้อยละ 20.20 (ME = 0.2020)

บทที่ 6

สรุป

งานวิจัยนี้มีวัตถุประสงค์ในวิเคราะห์หาความเสี่ยงของการเกิดความคลาดเคลื่อนของผลการวิจัยเชิงปริมาณด้านวิชาสังคมวิทยา ที่เป็นผลมาจากการละเมิดข้อตกลงในระเบียบวิธีการวิจัยเชิงปริมาณ

การนำเสนอในบทสรุป ผูวิจัยแบงนำเสนอเปน 3 สวน คือ สรุปผล อภิปรายผล และข้อเสนอแนะ โดยมีรายละเอียดดังต่อไปนี้

สรุปผลการวิจัย

ความคลาดเคลื่อนในการวิจัยมีความเกี่ยวข้องตั้งแต่เริ่มทำการวิจัยจนถึงการสรุปผลการวิจัย วิทยานิพนธ์และดุษฎีนิพนธ์ (งานวิจัย) ของผู้จบหลักสูตรระดับบัณฑิตศึกษา มีการเผยแพร่และให้บริการค้นคว้าสำหรับนำไปใช้ในการอ้างอิงทั้งเชิงวิชาการและการบริหาร แต่อาจมีการละเมิดข้อตกลงเบื้องต้น (basic assumption) ในขั้นการออกแบบการวิจัยและการวิเคราะห์ข้อมูล และทำให้เกิดความเสี่ยงในการเกิดความคลาดเคลื่อนชนิดที่ 1 (type I error) กับความคลาดเคลื่อนชนิดที่ 2 (type II error)

จากการรวบรวมข้อมูลทุติยภูมิ (secondary data) ที่เป็นงานวิจัยสาขาสังคมวิทยา ที่เผยแพร่อยู่ในฐานข้อมูลโครงการเครือข่ายห้องสมุดในประเทศ ของสำนักงานคณะกรรมการการอุดมศึกษา กระทรวงศึกษาธิการ ที่แล้วเสร็จระหว่างปี พ.ศ. 2512-2560 จำนวน 1,171 เรื่อง คัดเลือกเฉพาะงานวิจัยที่ใช้วิธีการวิจัยเชิงปริมาณเพียงวิธีการเดียวเท่านั้น ได้จำนวน 573 เรื่อง กำหนดขนาดกลุ่มตัวอย่าง (sample size) ด้วยสูตรของสูตรทาโร่ ยามาเน่ (Taro Yamane) ได้จำนวนตัวอย่าง 236 ตัวอย่าง และสุ่มตัวอย่างแบบอาศัยความน่าจะเป็นแบบง่าย (simple random sampling) โดยใช้คำสั่ง rand() สร้างเลขสุ่ม (random number) ในการเลือกตัวอย่าง

งานวิจัยที่เป็นกลุ่มตัวอย่างมากที่สุด คือ สาขาวิชาสังคมวิทยาประยุกต์ รองลงมา คือ สาขาสังคมวิทยา สาขาสังคมวิทยาและมานุษยวิทยา และน้อยที่สุด คือ สาขาสังคมวิทยาการพัฒนา และส่วนใหญ่เป็นงานวิจัยระดับปริญญาโทมากกว่าระดับปริญญาเอก และจากการวิเคราะห์วิธีการดำเนินการวิจัยได้ผลสรุปดังนี้

1. งานวิจัยส่วนใหญ่มีการตั้งสมมติฐานมากกว่าไม่มีการตั้งสมมติฐาน

2. งานวิจัยส่วนใหญ่มีการตรวจสอบเครื่องมือมากกว่าไม่มีการตรวจสอบเครื่องมือ

3. งานวิจัยส่วนใหญ่ใช้ระดับการวัดตัวแปรเป็นแบบค่าแบ่งกลุ่มมากกว่าค่าต่อเนื่อง

4. งานวิจัยส่วนใหญ่ประชากรที่นำมาใช้ในการศึกษาเป็นประชากรที่ทราบจำนวนมากกว่าไม่ทราบจำนวน

5. งานวิจัยส่วนใหญ่ใช้วิธีการกำหนดขนาดตัวอย่างแบบไม่ใช้วิธีการทางสถิติมากกว่าใช้วิธีการทางสถิติ

6. งานวิจัยส่วนใหญ่เลือกตัวอย่างใช้หลักความน่าจะเป็นมากกว่าไม่ใช้หลักความน่าจะเป็น

7. งานวิจัยเกือบทั้งหมดไม่มีการตรวจสอบข้อมูลเบื้องต้นก่อนการวิเคราะห์ข้อมูลด้วยสถิติ

การวิเคราะห์ความเสี่ยงของการเกิดความคลาดเคลื่อนในเบื้องต้น เป็นการพิจารณาจากสัดส่วน (ratio) ของตัวแปรที่เป็นองค์ประกอบภายในกรอบแนวคิดการวิจัยจำนวน 8 ตัวแปร พบว่า ตัวแปรที่น่าจะมีผลทำให้เกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากที่สุด คือ การตรวจสอบข้อมูล รองลงมา คือ สมมติฐาน ขนาดตัวอย่าง ระดับการวัด สถิติ การเลือกตัวอย่าง เครื่องมือวัด และน้อยที่สุด คือ ประชากร

จากการวิเคราะห์ข้อมูลด้วยวิธีการวิเคราะห์จัดกลุ่ม (Cluster Analysis: CA) แบ่งงานวิจัยที่มีรูปแบบวิธีการดำเนินการวิจัยที่เหมือนกันได้จำนวน 30 กลุ่ม มีกลุ่มงานวิจัยที่ไม่มีการทดสอบสมมติฐาน 4 กลุ่ม จึงตัดออกจากการวิเคราะห์ เหลือกลุ่มงานวิจัยที่มีการทดสอบสมมติฐาน 26 กลุ่ม เหลืองานวิจัยที่เป็นตัวอย่างจำนวน 225 ตัวอย่าง (เรื่อง) โดยพบว่า ตัวอย่างที่มีความเสี่ยงเกิดความคลาดเคลื่อนมากที่สุด 5 อันดับมีทั้งงานวิจัยที่ใช้สถิติเชิงอ้างอิงแบบมีพารามิเตอร์และไม่มีพารามิเตอร์

เมื่อวิเคราะห์ภาพรวมโอกาสของการเกิดความเสี่ยง พบว่า มีงานวิจัยที่เป็นกลุ่มตัวอย่างมีความเสี่ยงมากกว่าไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อน แต่จากการทดสอบสัดส่วนเชิงสถิติ พบว่า มีสัดส่วนระหว่างมีความเสี่ยงและไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อนไม่แตกต่างกัน

วิทยานิพนธ์ที่เป็นตัวอย่างในกลุ่มที่มีความเสี่ยงในการเกิดความคลาดเคลื่อน มีการละเมิดข้อตกลงมากที่สุด คือ การตรวจสอบข้อมูล รองลงมา คือ ระดับการวัด การเลือกตัวอย่าง ขนาดตัวอย่าง ประชากร และเครื่องมือวัด

การวิเคราะห์การทำนายโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยด้วยวิธีการวิเคราะห์การถดถอยแบบพหุโลจิสติก (Multiple Logistic Regression: MLR) โดยใช้ตัวแปรหุ่นตามกรอบแนวคิดของการวิจัย คือ เครื่องมือวัด ระดับการวัด ประชากร การกำหนดขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล ได้ผลสรุปดังนี้

1. วิธีนำตัวแปรทำนายเข้าวิเคราะห์ในแบบจำลองพร้อมกันทั้งหมด (Enter Method) พบว่า สมการที่ประกอบด้วยตัวแปรทำนาย 6 ตัวแปร คือ เครื่องมือวัด ระดับการวัด ประชากร ขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล โดยตัวแปรทำนายสามารถทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยได้ประมาณร้อยละ 37.7 (NR^2 = 0.377) และแบบจำลองประสิทธิภาพในการทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยประมาณร้อยละ 74.2 ตัวแปรทำนายไม่มีผลกระทบร่วมกันต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ (Sig. = 0.999) แต่มีตัวแปรทำนายที่มีผลกระทบต่อความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยอย่างมีนัยสำคัญทางสถิติ 2 ตัวแปร คือ เครื่องมือวัด (Sig. = 0.000) และขนาดตัวอย่าง (Sig. = 0.002)

2. วิธีการคัดเลือกตัวแปรทำนายออกจากแบบจำลองอย่างเป็นขั้นตอนโดยใช้ค่าสัดส่วนความควรจะเป็น (Backward Stepwise: Likelihood Ratio) ได้สมการที่ประกอบด้วยตัวแปรทำนาย 3 ตัวแปร คือ เครื่องมือวัด ประชากร และขนาดตัวอย่าง โดยตัวแปรทำนายสามารถทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยได้ประมาณร้อยละ 35.5 (NR^2 = 0.355) และแบบจำลองมีประสิทธิภาพในการทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยโดยรวมได้ถูกต้องประมาณร้อยละ 71.6 ตัวแปรทำนายมีผลกระทบร่วมกันต่อความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยอย่างมีนัยสำคัญทางสถิติ (Sig. = 0.000) และสามารถทำนายได้ว่างานวิจัยที่มีการละเมิดข้อตกลง คือ ไม่มีการตรวจสอบเครื่องมือวัด ใช้ประชากรที่ไม่ทราบจำนวน และไม่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง มีความเสี่ยงในการเกิดความคลาดเคลื่อนในงานวิจัย (Prob. = 0.85)

3. วิธีคัดเลือกตัวแปรทำนายเข้าแบบจำลองอย่างเป็นขั้นตอน (Forward Stepwise) ได้สมการที่ประกอบด้วยตัวแปรทำนาย 2 ตัวแปร คือ เครื่องมือวัด และขนาดตัวอย่าง โดยตัวแปรทำนายสามารถทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยได้ประมาณร้อยละ 34.2 (NR^2 = 0.342) และแบบจำลอง

มีประสิทธิภาพในการทำนายความถูกต้องในการแบ่งกลุ่มความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัย โดยรวมได้ถูกต้องประมาณร้อยละ 68.9 ตัวแปรทำนายมีผลกระทบร่วมกันต่อความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยอย่างมีนัยสำคัญทางสถิติ (Sig. = 0.000) และสามารถทำนายได้ว่างานวิจัยที่มีการละเมิดข้อตกลง คือ ไม่มีการตรวจสอบเครื่องมือวัดและไม่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง มีความเสี่ยงในการเกิดความคลาดเคลื่อนในงานวิจัย (Prob. = 0.93)

เมื่อพิจารณาจากค่าความสัมพันธ์ระหว่างโอกาสเกิดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัย (Odds Ratio) สรุปได้ดังนี้

1. งานวิจัยที่ละเมิดข้อตกลงโดยไม่มีการตรวจสอบเครื่องมือวัด (เครื่องมือวัด = 1) มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงโดยมีการทดสอบเครื่องมือวัด (เครื่องมือวัด = 0) ประมาณ 14 เท่า (OR = 14.059)

2. งานวิจัยที่ละเมิดข้อตกลงโดยไม่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง (ขนาดตัวอย่าง = 1) มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงโดยใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง (ขนาดตัวอย่าง = 0) ประมาณ 2 เท่า (OR = 2.322)

จากการวิเคราะห์ผลกระทบส่วนเพิ่ม (Marginal Effect) ของตัวแปรทำนายได้ขนาดผลกระทบส่วนเพิ่มของตัวแปรทำนายที่มีความสามารถในการทำนายความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยอย่างมีนัยสำคัญทางสถิติ 2 ตัวแปร คือ เครื่องมือวัด (Sig. = 0.000) และขนาดตัวอย่าง (Sig. = 0.006) สรุปได้ดังนี้

1. งานวิจัยที่ละเมิดข้อตกลงโดยไม่มีการทดสอบเครื่องมือวัด (เครื่องมือวัด = 1) มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงโดยมีการทดสอบเครื่องมือวัด (เครื่องมือวัด = 0) ร้อยละ 49.93 (ME = 0.4993)

2. งานวิจัยที่ละเมิดข้อตกลงโดยไม่ใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง (ขนาดตัวอย่าง = 1) มีความเสี่ยงในการเกิดความคลาดเคลื่อนในการวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงโดยใช้วิธีการทางสถิติในการคำนวณขนาดตัวอย่าง (ขนาดตัวอย่าง = 0) ร้อยละ 20.20 (ME = 0.2020)

คำตอบของปัญหาการวิจัย

การวิจัยเชิงปริมาณสาขาสังคมวิทยาของมหาวิทยาลัยในประเทศไทยที่เผยแพร่อยู่ในฐานข้อมูลของโครงการเครือข่ายห้องสมุดในประเทศไทยมีการออกแบบวิธีการดำเนินการวิจัยที่เหมือนกันจำนวน 30 กลุ่ม

งานวิจัยที่เป็นกลุ่มตัวอย่างมีความเสี่ยงมากกว่าไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อน แต่จากการทดสอบสัดส่วนเชิงสถิติ พบว่า มีสัดส่วนระหว่างมีความเสี่ยงและไม่มีความเสี่ยงในการเกิดความคลาดเคลื่อนไม่แตกต่างกัน

งานวิจัยที่เป็นตัวอย่างในกลุ่มที่มีความเสี่ยงในการเกิดความคลาดเคลื่อน มีการละเมิดข้อตกลงมากที่สุด คือ การตรวจสอบข้อมูล รองลงมา คือ ระดับการวัด การเลือกตัวอย่าง ขนาดตัวอย่าง ประชากร และเครื่องมือวัด

การวิเคราะห์การทำนายโอกาสที่จะมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยโดยใช้ตัวแปรเครื่องมือวัด ระดับการวัด ประชากร การกำหนดขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล พบว่า งานวิจัยที่มีการละเมิดข้อตกลงด้านเครื่องมือวัดมีความเสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงด้านเครื่องมือวัดประมาณ 14 เท่า หรือประมาณร้อยละ 49.93 และงานวิจัยที่มีการละเมิดข้อตกลงด้านขนาดตัวอย่างมีความสี่ยงของการเกิดความคลาดเคลื่อนในงานวิจัยมากกว่างานวิจัยที่ไม่ละเมิดข้อตกลงด้านขนาดตัวอย่างประมาณ 2 เท่า หรือประมาณร้อยละ 20.20

อภิปรายผล

การวิจัยความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยเชิงปริมาณสาขาสังคมวิทยา มีประเด็นสำคัญของผลการวิจัยที่สามารถอธิบายนำมาอภิปรายโดยมีรายละเอียดดังนี้ต่อไปนี้

จากผลการวิจัยพบว่า วิทยานิพนธ์และดุษฎีนิพนธ์ส่วนใหญ่มีการตั้งสมมติฐาน มีการตรวจสอบเครื่องมือ มีระดับการวัดเป็นแบบค่าแบ่งกลุ่ม ใช้ประชากรแบบทราบจำนวน กำหนดขนาดตัวอย่างแบบไม่ใช้วิธีการทางสถิติ เลือกตัวอย่างใช้หลักความน่าจะเป็น แต่เกือบทั้งหมดไม่มีการตรวจสอบข้อมูลก่อนการวิเคราะห์ข้อมูลด้วยสถิติ ซึ่งสอดคล้องกับที่แกรี่ ที เฮนรี่ (Henry, 1990: 34-46) อธิบายไว้ว่า ความคลาด

เคลื่อนโดยรวม (total error) มีองค์ประกอบสาเหตุมาจากความคลาดเคลื่อนที่ไม่ได้เกิดจากการเลือกตัวอย่าง (nonsampling bias) ได้แก่ กรอบและรายการของตัวอย่าง การไม่ตอบคำถาม ความคลาดเคลื่อนจากการวัด และความลำเอียงที่เกิดจากการเลือกตัวอย่าง (sampling bias) ได้แก่ ความลำเอียงจากการเลือกตัวอย่าง ความลำเอียงจากการประมาณการ และความแปรปรวนจากการเลือกตัวอย่าง (sampling variability) ได้แก่ ขนาดตัวอย่าง และความเป็นหนึ่งเดียวของตัวอย่าง

งานวิจัยที่มีสมมติฐาน มีความเสี่ยงเกิดความคลาดเคลื่อนมากว่างานวิจัยที่ไม่มีสมมติฐาน เพราะในกระบวนการทดสอบสมมติฐานด้วยสถิติมีโอกาสเกิดความคลาดเคลื่อนจากละเมิดข้อตกลงในการออกแบบเครื่องมือ การเลือกตัวอย่าง การตรวจสอบข้อมูล และการเลือกใช้สถิติ ซึ่งสอดคล้องกับผลการวิจัยที่พบว่างานวิจัยส่วนใหญ่ไม่มีการตรวจสอบข้อมูลจึงมีโอกาสที่จะเกิดความคลาดเคลื่อนในการวิเคราะห์ด้วยสถิติตามมา เช่น ให้ผลการวิเคราะห์ผิดพลาด (bias) และค่าสถิติไม่มีอำนาจในการทำนายที่ดี (statistical power) และยังพบว่า งานวิจัยส่วนใหญ่คำนวณขนาดตัวอย่างโดยไม่ใช้วิธีการทางสถิติ ซึ่งขนาดตัวอย่างมีผลต่อความคลาดเคลื่อนชนิดที่ 1 และความคลาดเคลื่อนชนิดที่ 2 กล่าวคือ ขนาดตัวอย่างที่น้อยหรือไม่เพียงพออาจทำให้เกิดความคลาดเคลื่อนชนิดที่ 1 แต่หากเพิ่มขนาดตัวอย่างหรือมีจำนวนขนาดตัวอย่างมากเกินไปอาจทำให้เกิดความคลาดเคลื่อนชนิดที่ 2 มากขึ้น

งานวิจัยที่มีการทดสอบสมติฐานด้วยสถิติแบบมีพารามิเตอร์ มีการละเมิดของตกลงในกระบวนการดำเนินการวิจัยในขั้นตอนใดขั้นตอนหนึ่งมีโอกาสทำให้เกิดความคลาดเคลื่อนในการยอมรับหรือปฏิเสธสมมติฐาน คือ ไม่มีการทดสอบเครื่องมือ ใช้ประชากรแบบไม่ทราบจำนวน กำหนดขนาดตัวอย่างแบบไม่ใช้วิธีการทางสถิติ เลือกตัวอย่างโดยไม่ใช้หลักความน่าจะเป็น ข้อมูลมีระดับการวัดไม่ใช่ค่าแบ่งกลุ่ม และไม่มีการตรวจสอบข้อมูลตามข้อตกลงเบื้องต้นของการใช้สถิติแบบมีพารามิเตอร์ ส่วนงานวิจัยที่มีการทดสอบสมมติฐานด้วยสถิติแบบไม่มีพารามิเตอร์ แม้ว่าจะเป็นสถิติทดสอบที่เป็นอิสระจากข้อตกลง (assumption-free tests) แต่พบกว่า ไม่มีการทดสอบเครื่องมือ จึงเป็นการละเมิดข้อตกลงในกระบวนการดำเนินการวิจัยในขั้นตอนใดขั้นหนึ่งเหมือนกัน

จากการจำลองตัวแบบและวิเคราะห์ด้วยการวิเคราะห์ถดถอยแบบพหุโลจิสติกได้ตัวแปรที่อยู่ในกลุ่มความคลาดเคลื่อนจากเครื่องมือวัด คือ การตรวจสอบเครื่องมือวัด และตัวแปรที่อยู่ในกลุ่มความคลาดเคลื่อน

จากการสุ่มตัวอย่าง คือ การคำนวณขนาดตัวอย่าง ซึ่งสอดคล้องกับที่ พอล พี. บีเมอร์ (Biemer & others, 1991: 487) อธิบายไว้ว่า ความคลาดเคลื่อนของการวิจัยและแบ่งได้เป็น 2 กลุ่มหลักตามทฤษฎีของความคลาดเคลื่อน ดังนี้

1. กลุ่มที่เชื่อว่าความคลาดเคลื่อนเกิดมาจากการเลือกตัวอย่าง เน้นการอธิบายความแปรปรวน-ความแปรปรวนร่วม (variance-covariance) ในเชิงโครงสร้างของประชากรทั้งหมด โดยมีสมมติฐานว่า ความคลาดเคลื่อนเกิดมาจากข้อมูลหรือคำตอบที่ได้มาจากการเลือกตัวอย่างแบบสุ่มทั้งจากประชากรที่ทราบจำนวน (finite population) และประชากรที่ไม่ทราบจำนวน (infinite population)

2. กลุ่มที่เชื่อว่าความคลาดเคลื่อนเกิดมาจากเครื่องมือวัด เน้นศึกษาความสัมพันธ์ของคำตอบที่มีความหลากหลายที่ได้มาจากตัวอย่างทั้งหมด และประเมินความถูกต้องหรือความคลาดเคลื่อนของคำตอบที่ได้มาจากแต่ละบุคคล

อย่างไรก็ตามความคลาดเคลื่อนในงานวิจัยไม่ได้เกิดจาก 2 ความเชื่อดังกล่าวเท่านั้น เพราะเฮอร์แมน เจ. ซี. บีเรนด์เซ่น (Berendsen, 2011: 18-19) ได้อธิบายไว้ว่า ความคลาดเคลื่อนอาจเกิดได้จากความคลาดเคลื่อนเชิงบุคคล (gross error หรือ human error) ที่เกิดจากความเข้าใจผิดในการสร้างและใช้เครื่องมือวัดของนักวิจัย ได้แก่ ความไม่ตั้งใจจากการไม่ระมัดระวังในการตรวจสอบและตรวจซ้ำ ไม่รู้เท่าทันจากการละเลยหรือไม่เอาใจใส่ และจงใจ (intended) เลือกข้อมูลแบบเจาะจง

ปัจจัยที่มีผลต่อการเกิดความความคลาดเคลื่อนในการวิจัยเชิงปริมาณโดยรวมเกิดขึ้นได้ในทุกขั้นตอนในการทำวิจัย โดยไทสัน โฮล์มส์ (Holmes, 2004) อธิบายไว้ว่า มีความคลาดเคลื่อนเชิงสถิติ 10 ประเภท คือ ความลำเอียงในการเลือกตัวอย่าง ความไม่แม่นยำในการเลือกตัวอย่าง ความลำเอียงในการวัด ความไม่แม่นยำในการวัด ความลำเอียงในการประมาณการ ความไม่แม่นยำในการประมาณการ ความลำเอียงในการทดสอบสมมติฐาน ความไม่แม่นยำในการทดสอบสมมติฐาน ความลำเอียงในการรายงาน และความไม่แม่นยำในการรายงาน

เป้าหมายที่สำคัญของการใช้สถิติ คือ การนำเอาข้อมูลหรือสารสนเทศจากตัวอย่าง (sample) อ้างอิง (inference) ไปสู่การอธิบายหรือสรุปประชากร (population) (Scheaffer, Mendenhall III, & Ott, 2006: 48-50) การละเมิด (violation) ข้อตกลงบางครั้งอาจไม่มีผลต่อข้อสรุปของการวิจัยแต่อย่างไร แต่บางครั้งอาจ

มีผลต่อการแปลความหมายของผลการวิจัยเป็นอย่างมาก และไม่น่าที่จะทำให้ได้ผลการวิจัยที่ถูกต้องและได้ข้อสรุปที่น่าเชื่อถือ (Ghasemi & Zahediasl, 2012: 486-489) เกิดความคลาดเคลื่อนหรือความผิดพลาดแบบที่ 1 (type I (alpha) error) และความคลาดเคลื่อนหรือความผิดพลาดแบบที่ 2 (type II (beta) error) หรือทำให้การประมาณการค่านัยสำคัญ (significance) หรือขนาดอิทธิพล (effect size) ต่ำกว่าหรือมากกว่า แต่เรากลับยัดเยียดความถูกต้องไว้ในผลลัพธ์ ข้อสรุป และการยืนยันผลของการวิจัย โดยไม่ได้ทำการทดสอบตามหลักสถิติว่า เป็นไปตามข้อตกลงเบื้องต้นหรือไม่ (Osborne & Waters, 2002)

คนส่วนใหญ่คิดว่าสถิติเป็นเรื่องของพีชคณิต (algebra) และการคำนวณ (computation) ซึ่งเป็นมุมมองที่แคบและใช้ทดสอบนัยสำคัญ (significance testing) ในทางที่ผิด เช่น ใช้มากเกินไป ใช้น้อยเกินไป ใช้ผิด แปลความหมายผิด และใช้เพื่อให้น่าเชื่อถือ (Rigby, 1998) ซึ่งสอดคล้องกับที่แฟรงค์ ชมิดท์และจอห์น ฮันเตอร์ (Schmidt & Hunter, 1989, 1996 cited by Thye, 2000) กล่าวไว้ว่า บางกรณีมีการใช้สูตรผิดหรือบางกรณีใช้สูตรถูก แต่ตีความหมายผิด

เชน อาร์ ทาย (Thye, 2000) จำแนกสาเหตุของการเกิดความคลาดเคลื่อนในการวัด (measurement error) เป็น 2 ชนิด คือ ความคงที่ (constant) และความสุ่ม (random) โดยความคงที่ของความคลาดเคลื่อนในการวัดเป็นผลที่เกิดจากความไม่สมบูรณ์ของความเที่ยงตรงเชิงโครงสร้าง (construct validity) ในการสังเกตแต่ละครั้งด้วยวิธีการเดียวกันในการวัดแต่ละครั้ง และความสุ่มของความคลาดเคลื่อนในการวัดเป็นผลที่เกิดมาจากระดับความสุ่มในการสังเกตแต่ละครั้งมีความไม่น่าเชื่อถือ (unreliability) ดังนั้นความคลาดเคลื่อนทั้ง 2 ชนิด จึงมีความสำคัญต่อการพัฒนาทฤษฎีที่เป็นสากล (law-like theory)

ความคลาดเคลื่อนที่เกิดขึ้นไม่สามารถกำจัดออกไปได้จากการวิจัย แต่สามารถลดผลกระทบที่เกิดขึ้นได้ เช่น การสร้างแบบสอบถาม ให้มีคุณภาพและชัดเจน พยายามติดต่อและขอคำตอบใหม่ หากไม่ได้รับคำตอบหรือปฏิเสธไม่ให้คำตอบ และการตรวจสอบข้อมูล ให้ถูกต้องและสมบูรณ์ (Groves cited by Scheaffer, Mendenhall III, & Ott, 2006: 18-25; Thompson, 2012: 5)

อาจกล่าวได้ว่า ความคลาดเคลื่อนของงานวิจัยไม่ใช่เรื่องใหม่ เพราะมีหลักฐานที่แสดงให้เห็นว่า ในอดีตมีความสนใจการเกิดความคลาดเคลื่อนในงานวิจัยอยู่ เช่น จากบทความเรื่อง “ความคลาดเคลื่อนในการสำรวจ (On Errors in Surveys)” ของวิลเลียม เอ็ดเวิดส์ เดมมิ่ง (William Edwards Deming) น่าจะเป็น

หลักฐานเชิงประจักษ์ที่บอกได้ว่า การศึกษาเกี่ยวกับความคลาดเคลื่อนในการวิจัยด้านสังคมศาสตร์มีจุดเริ่มต้นในปี ค.ศ. 1914 จากข้อมูลการสัมภาษณ์เพื่อศึกษาด้านสังคมของสำนักงานสถิติแห่งชาติของสหรัฐอเมริกา (United State Census Bureau) ที่ปรากฏอยู่ในงานของสจ๊วต เอ ไรซ์ (Struart A, Rice) เรื่อง “ความลำเอียงเป็นโรคที่ติดต่อกันได้ในการสัมภาษณ์ (Contagious bias in the interview)” ที่พบว่า ผู้ให้ข้อมูลให้เหตุผลของความยากจนแตกต่างกันไปตามการปรุงแต่งของผู้สัมภาษณ์ (Rice 1929; Deming 1944)

ในปี ค.ศ. 1979 ฮูเบิร์ต บลาล็อค (Hubert Blalock) นายกสมาคมสังคมวิทยาอเมริกัน (American Sociological Association) ได้กล่าวว่า การวัด (measurement) เป็นปัญหาสำคัญและร้ายแรงที่กำลังเผชิญอยู่ในการเรียนการสอนด้านสังคมวิทยา แต่หลังจากนั้นเป็นต้นมามีการเปลี่ยนแปลงไม่มากนัก ทุกวันนี้นักวิจัยยังคงพบกับปัญหาในการวัดและการสร้างกรอบแนวคิด (conceptualization) ที่ไม่แตกต่างไปจากที่บลาล็อคกังวลไว้ตั้งแต่ในอดีต ความคลาดเคลื่อนเชิงสถิติ (statistical errors) ในงานวิจัยเชิงปริมาณของนักสังคมศาสตร์คงเป็นเรื่องซ้ำซากเช่นเดียวกับงานวิจัยเชิงวิทยาศาสตร์ของนักวิทยาศาสตร์ที่เกิดขึ้นมาอย่างยาวนาน เพราะการขาดความรู้และความเข้าใจเกี่ยวกับพื้นฐานของสถิติ (basic of statistics) ขาดการสอนที่ดีโดยนักสถิติ ขาดความสนใจและความเอาใจใส่หลังจากเรียนจบไปแล้ว จึงทำให้มีทักษะด้านสถิติที่ไม่เพียงพอในการทำวิจัยที่ต้องใช้สถิติ (Rigby 1998; Thye 2000) และมีการประมาณการว่า ร้อยละ 50 ของบทความที่เป็นเชิงวิทยาศาสตร์ที่ตีพิมพ์ มีอย่างน้อย 1 บทความที่มีความคลาดเคลื่อนเชิงสถิติ (Curran-Everett & Benos 2004)

การเกิดความคลาดเคลื่อนในงานวิจัยยังคงมีอยู่ต่อไป เพราะมีหลายปัจจัยเข้ามาเกี่ยวข้องไม่ใช่เฉพาะระบบเบียบวิธีและสถิติที่ในการวิเคราะห์เท่านั้น เช่น การทบทวนวรรณกรรม ความนิยมชมชอบในแนวคิดและทฤษฎี รวมถึงปรากฏการณ์ทางสังคมที่มีการเปลี่ยนแปลงอย่างรวดเร็วจนไม่สามารถทำนายและสรุปได้ด้วยการวิจัยที่ต้องใช้เวลาในการเดินตามระเบียบวิธีที่เคร่งครัด แต่หากไม่มีหลักการหรือแนวปฏิบัติเลย การศึกษาเหตุการณ์อุบัติใหม่ทางสังคม (social emergence) จะมีความคลาดเคลื่อนเกิดขึ้นอีกมากมาย

ข้อเสนอแนะ

ข้อเสนอแนะเชิงวิชาการ

การวิจัยมีความสำคัญต่อการนำเอาผลการวิจัยไปใช้และพัฒนาต่อยอดทางวิชาการ ดังนั้นในการทำวิทยานิพนธ์หรือดุษฎีนิพนธ์ รวมถึงงานวิจัยประเภทอื่นๆ ต้องตระหนักและเข้มงวดในการดำเนินการวิจัยให้เป็นไปตามข้อตกลงและระเบียบวิธีการวิจัยอย่างเคร่งครัด เพื่อป้องกันการเกิดความคลาดเคลื่อน โดยเริ่มจากการหาความรู้พื้นฐานและต่อยอดความรู้ด้านการวิจัยอย่างเป็นระบบ ผู้วิจัยที่ต้องการใช้วิธีการวิจัยเชิงปริมาณต้องมีการศึกษาทั้งด้านระเบียบวิธีการวิจัยเชิงปริมาณและการวิเคราะห์ข้อด้วยสถิติให้เกิดความเข้าใจอย่างท่องแท้ก่อนปฏิบัติการวิจัยด้วยวิธีการดังกล่าว การนำงานวิจัยเชิงปริมาณเข้าในระบบฐานข้อมูลเพื่อเผยแพร่ในวงกว้างควรกำหนดให้เจ้าของงานวิจัยสร้างเอกสารแสดงข้อมูล (data sheet) แสดงรายการการทำตามข้อตกลงเบื้องต้นให้ชัดเจน ซึ่งจะเป็นประโยชน์ต่อผู้นำผลงานวิจัยไปใช้และสามารถนำไปใช้พัฒนาเป็นดัชนีแสดงระดับความคลาดเคลื่อนของงานวิจัยได้ด้วย

ข้อเสนอแนะเชิงการนำไปใช้

จากผลการวิจัยแสดงให้เห็นว่า มีการละเมิดข้อตกลงในแต่ละขั้นตอนการดำเนินการวิจัยอย่างเห็นได้ชัด การเลือกใช้สถิติแบบมีพารามิเตอร์ไม่สามารถละเมิดข้อตกลงข้อใดข้อหนึ่งได้ เพราะข้อตกลงแต่ละข้อมีความเชื่อมโยงและเกี่ยวข้องกันตลอดกระบวนการวิจัย แม้ว่าสถิติแบบไม่มีพารามิเตอร์จะเป็นอิสระจากข้อตกลง (assumption-free tests) แต่หากไม่มีการตรวจสอบเครื่องมือวัดก็สามารถทำให้เกิดความคลาดเคลื่อนได้ เพราะเครื่องมือวัดเน้นศึกษาความสัมพันธ์ของคำตอบที่มีความหลากหลายที่ได้มาจากตัวอย่างทั้งหมด และประเมินความถูกต้องหรือความคลาดเคลื่อนของคำตอบที่ได้มาจากแต่ละบุคคล ดังนั้นผู้นำผลงานวิจัยไปใช้ประโยชน์ควรมีการตรวจสอบว่างานวิจัยมีการละเมิดข้อตกลงเบื้องต้นหรือไม่ เพื่อให้มั่นใจว่าผลงานวิจัยไม่มีความคลาดเคลื่อน

ข้อเสนอแนะเชิงวิจัย

การวิเคราะห์ความเสี่ยงการเกิดความคลาดเคลื่อนในการวิจัย เป็นการเริ่มให้ความสนใจถึงความถูกต้องของผลการวิจัยจากประสบการณ์ที่ได้พบว่า ข้อมูลที่ได้จากการวิจัยบางเรื่องเมื่อเลือกวิเคราะห์ด้วยสถิติ

แบบมีพารามิเตอร์ไม่พบนัยสำคัญทางสถิติ และเมื่อนำไปวิเคราะห์ด้วยสถิติแบบไม่มีพารามิเตอร์พบนัยสำคัญทางสถิติ งานวิจัยนี้ก็มีโอกาสเกิดความคลาดเคลื่อนได้เช่นกัน แม้ว่าจะระมัดระวังเกี่ยวกับการละเมิดข้อตกลงตลอด เพราะมีหลายสาเหตุที่อาจทำให้ได้ข้อสรุปที่ไม่ถูกต้อง เช่น งานวิจัยบางเล่มไม่ได้เขียนสาระสำคัญเกี่ยวกับเครื่องมือวัด ระดับการวัด ที่มาของประชากร การกำหนดขนาดตัวอย่าง การเลือกตัวอย่าง และการตรวจสอบข้อมูล ทำให้การรวบรวมข้อมูลอาจไม่ตรงกับข้อเท็จจริงก็เป็นได้ ดังนั้นอาจต้องมีการศึกษาซ้ำเพื่อยืนยันผลการวิจัยดังกล่าว

ปัจจุบัน วิทยานิพนธ์และดุษฎีนิพนธ์ข้อมูลโครงการเครือข่ายห้องสมุดในประเทศ ของสำนักงานคณะกรรมการการอุดมศึกษา มีการจัดเก็บบางส่วนเป็นไฟล์ข้อความ เช่น บทคัดย่อ ที่คอมพิวเตอร์สามารถสามารถอ่านได้ (machine readable data) แต่อาจมีรายละเอียดเกี่ยวกับข้อตกลงเบื้องต้นไม่พอที่จะสกัดข้อมูล (extract data) ออกมาและใช้การเรียนรู้โดยเครื่อง (machine learning) วิเคราะห์หาความเสี่ยงการเกิดความคลาดเคลื่อนในงานวิจัยได้ ดังนั้นจึงควรกำหนดให้มีการแปลงเอกสารงานวิจัยให้เป็นไฟล์ที่คอมพิวเตอร์สามารถอ่านและนำไปประมวลผลได้ เช่น CSV, RDF, XML และ JSON

เอกสารอ้างอิง

ภาษาไทย

มหาวิทยาลัยเชียงใหม่. (2530). *สังคมวิทยาและมานุษยวิทยาไทย*. เชียงใหม่: ภาควิชาสังคมวิทยาและมานุษยวิทยา คณะสังคมศาสตร์ มหาวิทยาลัยเชียงใหม่.

เรวัต แสงสุริยงค์. (2541). *เอกสารประกอบการสอนวิชา 225101 สังคมวิทยาเบื้องต้น*. ชลบุรี: ภาควิชาสังคมวิทยา มหาวิทยาลัยบูรพา.

สัญญา สัญญาวิวัฒน์ (2525). *วิวัฒนาการของสังคมวิทยาในรอบ 200 ปี ของกรุงรัตนโกสินทร์*. กรุงเทพฯ: โครงการไทยศึกษา ฝ่ายวิชาการ จุฬาลงกรณ์มหาวิทยาลัย.

ภาษาอังกฤษ

Allison, P. D. (2002). *Missing Data*. Thousan Oaks, Calif: SAGE.

Aneshensel, C. S. (2013). *Theory-Based Data Analysis for the Social Sciences (2 ed.)*. Los Angeles: SAGE.

Arnab, R. (2017). *Survey Sampling Theory and Applications*. London: Academic Press.

Aven, T. (2011). *Quantitative risk assessment: the scientific platform*. Cambridge: Cambridge University Press.

Awang, Z., Affthanorhan, A. & Asri, M.A.M. (2015). Parametric and Non Parametric Approach in Structural Equation Modeling (SEM): The Application of Bootstrapping. *Modern Applied Science,* 9(9), 58-67. DOI: 10.5539/mas.v9n9p58

Babbie, E. (2013). *The Practice of Social Research (13 ed.).* Andover: Wadsworth Cengage Learning.

Bartlett, J. E., Kotrlik, J.W., & Higgins, C.C. (2001). Organizational Research: Determining Appropriate Sample Size in Survey Research. *Information Technology, Learning, and Performance Journal,* 19(1), 43-50.

Berendsen, H. J. C. (2011). *A Student's Guide to Data and Error Analysis*. Cambridge: Cambridge University Press.

Bevington, P. R. (1992). *Data reduction and error analysis for the physical sciences*. New York: McGraw-Hill.

Bevington, P. R., & Robinson, D.K. (1992). *Data reduction and error analysis for the physical sciences*. New York: McGraw-Hill.

Biemer, P. P. (1991). *Measurement errors in surveys*. New York: Wiley.

Blaikie, N. (2010). *Designing Social Research: The Logic of Anticipation (2 ed.)*. Cambridge, UK: Polity.

Blair, E. & Blair, J. (2015). *Applied Survey Sampling*. Los Angeles: SAG

Bollinger, C. R. (1998). Measurement Error in the Current Population Survey: A Nonparametric Look. *J Labor Econ,* 16(3), 576-594. DOI: 10.1086/209899

Bross, I. (1954). Misclassification in 2 X 2 Tables. *Biometrics,* 10(4), 478-486. DOI: 10.2307/3001619

Buonaccorsi, J. P. (2010). *Measurement Error: Models, Methods, and Applications*. Boca Raton: CRC Press.

Burdick, R. K. & Watrin, S. (2008). Confidence Intervals for the Coefficient of Variation when Random Effects are Selected from Finite Populations. *Quality Engineering*, 20(3), 261-268. DOI: 10.1080/08982110701733059

Burmeister, E., & Aitken, L. M. (2012). Sample size: How many is enough? *Australian Critical Care,* 25(4), 271-274. DOI: 10.1016/j.aucc.2012.07.002

Cochran, W. G. (1977). *Sampling Techniques*. Singapore: John Wiley & Sons.

Cohen, J. (1962). The Statistical Power of Abnormal-Social Psychological Research: A Review. *Journal of Abnormal and Social Psychology,* 65(3), 145-153. DOI: 10.1037/h0045186

Cornfield, J., Haenszel, W., Hammond, E. C., Lilienfeld, A. M., Shimkin, M. B., & Wynder, E. L.

(2009). Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions. *International Journal of Epidemiology,* 38(5), 1175-1191. DOI: 10.1093/ije/dyp289

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (4 ed.)*. Los Angeles: SAG.

Curran-Everett, D., & Benos, D. J. (2004). Guidelines for Reporting Statistics in Journals Published by the American Physiological Society. *Advances in Physiology Education,* 31(4), E189-E191. DOI: 10.1152/advan.00022.2007

Delice, A. (2010). The Sampling Issues in Quantitative Research. Education Sciences. *Theory & Practice,* 10(4), 2001-2018.

Deming, W. E. (1944). On Errors in Surveys. *American Sociological Review,* 9(4), 359-369. DOI: 10.2307/2085979

Drost, E. A. (2011). Validity and Reliability in Social Science Research. *Education Research and Perspectives*, 38 (1), 105-123. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ942587&site=eds-live

Evans, A. N. (2014). *Using Basic Statistics in the Behavioral and Social Sciences (5 ed.)*. Los Angeles: SAGE.

Fenton, N., Neil, M. (2019). *Risk assessment and decision analysis with bayesian networks*. Boca Raton, FL: CRC Press, Taylor & Francis Group.

Field, A. (2009). *Discovering statistics using SPSS (sex and drugs and rock 'n') roll*. Los Angeles: SAGE.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics: and sex and drugs and rock 'n' roll*. Los Angeles: SAGE.

Flynn, B. B., Sakakibara, S., Schroeder, R. G., Bates, K. A., & Flynn, E. J. (1990). Empirical Research Methods in Operations Management. *Journal of Operations Management,*

9(2), 250-284. DOI: 10.1016/0272-6963(90)90098-X

Garner, M., Wagner, C., & Kawulich, B. (2009). *Teaching Research Methods in the Social Sciences*. Farnham, Surrey, England: Burlington.

Garson, G. D. (2012). *Testing Statistical Assumptions*. North Carolina: Statistical Publishing Associates.

Ghasemi, A., & Zahediasl., S. (2012). Normality Tests for Statistical Analysis: A Guide for Non-statisticans. *International Journal of Endocrinology & Metabolism,* 10(2), 486-489. DOI: 10.5812/ijem.3505

Goldstein, R. (1989). Power and Sample Size via MS/PC-DOS Computers. *The American Statistician,* 43(4), 253-260. DOI: 10.2307/2685373

Hair, J. F., & Others (2006). *Multivariate Data Analysis*. Upper Saddle River, N.J.: Pearson/Prentice Hall.

Healey, J. F. (1999). *Statistics: A Tool for Social Research*. Belmont: Wadsworth.

Henry, G. T. (1990). *Practical Sampling*. London: SAGE.

Holmes, T. H. (2004). Ten Categories of Statistical Errors: A Guide for Research in Endocrinology and Metabolism. *Am J Physiol Endocrinol Metab,* 286(4), E495-501. DOI: 10.1152/ajpendo.00484.2003

Hopkin, P. (2017). *Fundamentals of risk management: understanding, evaluating and implementing effective risk management*. London: KoganPage.

Israel, G. D. (1992). *Determining Sample Size*. University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS. Retrieved from https://www.tarleton.edu/academicassessment/documents/Samplesize.pdf

Kennedy, J. E. (2015). Beware of Inferential Errors and Low Power with Bayesian Analyses: Power Analysis is Needed for Confirmatory Research. *Journal of Parapsychology,* 79(1), 53-64.

Kimberlin, C. L. & Winterstein, A. G. (2008). Validity and reliability of measurement Instruments used in research. *Am J Health Syst Pharm*, 65(23), 2276-2284. DOI: 10.2146/ajhp070364.

Krejcie, R. V., & Morgan, D. W. (2016). Determining Sample Size for Research Activities. *Educational and Psychological Measurement,* 30(3), 607-610. DOI: 10.1177/001316447003000308

Kumar, R. (2014). *Research Methodology: a step-by-step guide for beginners (4 ed.)*. Los Angeles: SAGE.

Lash, T. L., Fox, M. P., & Fink, A. K. (2009). *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York: Springer.

Lash, T. L., Fox, M. P., Cooney, D., Lu, Y., & Forshee, R. A. (2016). Quantitative Bias Analysis in Regulatory Settings. *Am J Public Health,* 106(7), 1227-1230. DOI: 10.2105/AJPH.2016.303199

Lehmann, E. L. (2005). *Testing statistical hypotheses*. New York: Springer.

Lenth, R. V. (2001). Some Sractical Guidelines for Effective Sample Size Determination. *American Statistician,* 55(3), 187-193. DOI: 10.1198/000313001317098149

Li, Q. (2018). *Using R for data analysis in Social Sciences: A Research Project-oriented Approach.* New York: Oxford University Press

Osborne, J. W. (2008). *Best practices in quantitative methods*. Thousand Oaks, Calif.: SAGE.

Osborne, J. W., & Waters, E. (2002). Four Assumptions of Multiple Regression that Researches Should Always Test. *Practical Assessment, Research & Evaluation*, 8(2), 1-5. Retrieved from https://pareonline.net/getvn.asp?v=8&n=2

Pritchard, C. L. (2010). *Risk management: concepts and guidance*. Arlington, VA: ESI International.

Rao, P. S. R. S. (2000). *Sampling Methodologies with Applications*. Boca Raton, Florida.

Rice, S. A. (1929). Contagious Bias in the Interview: A Methodological Note. *American Journal of Sociology,* 35(3), 420-423. DOI: 10.1086/215055

Rigby, A. S. (1998). Statistical Methods in Epidemiology: I. Statistical Errors in Hypothesis Testing. *Disabil Rehabil,* 20(4), 121-126. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9571378

Russell, J. W. (1992). *Introduction to Macrosociology*. Englewood Cliffs, N.J.: Prentice Hall.

Scheaffer, R. L., Mendenhall III, W., & Ott, R.L. (2006). *Elementary Survey Sampling*. Australia: ThomsonBrooks/Cole.

Schennach, S. M. (2012). *Measurement Error in Nonlinear Models: A Review*. DOI: 10.1017/CBO9781139060035.009

Sheskin, D. (2007). *Handbook of parametric and nonparametric statistical procedures*. Boca Raton: Chapman & Hall/CRC.

Singh, A. S., & Masuku., M. B. (2014). Sampling Techniques & Determination of Sample Size in Applied Statistics Research: An Overview. *International Journal of Economics, Commerce and Management,* 2(11), 1-22. Retrieved from http://ijecm.co.uk/wp-content/uploads/2014/11/21131.pdf

Singh, P., Singh, R., & Bouza, C. N. (2018). Effect of Measurement Error and Non-response on Estimation of Population Mean. *Investigation Operational,* (1), 108-120. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.534541787&site=eds-live&authtype=uid

Taherdoost, H. (2016). Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. *International Journal of Academic Research in Management (IJARM)*, 5(3), 28-36. DOI: 10.2139/ssrn.3205040

Téllez, A., García, C.H., & Corral-Verdugo V. (2015). Effect Size, *Confidence Intervals and Statistical Power in Psychological Research. Psychology in Russia: State of Art,* 8(3), 27-47. DOI: 10.11621/pir.2015.0303

Thomas, L. & Krebs, C. J. (1997). A review of statistical power analysis software. *Bulletin of the Ecological Society of America, 78(2),* 126-139. DOI: 10.2307/20168137

Thompson, S. K. (2012). *Sampling.* Hoboken, N.J.: Wiley.

Thye, S. R. (2000). Reliability in Experimental Sociology. *Social Forces,* 78(4), 1277-1309. DOI: 10.1093/sf/78.4.1277

Traub, R. E. (1994). *Reliability for the Social Sciences: Theory and Applications*. Thousand Oaks, Calif.: Sage.

Verma, J. P., G. & Abdel-Salam, A. G. (2019). *Testing Statistical Assumptions in Research*. Hoboken, N.J.: Wiley.

Wilkinson, D. E. (2000). *The Researcher's Toolkit: The Complete Guide to Practitioner Research*. London: Routledge Falmer.

Williams, M. (2016). *Key Concepts in the Philosophy of Social Research*. Los Angeles: SAGE.

Yamane, T. (1967). *Elementary Sampling Theory*. Englewood Cliffs: Prentice-Hall.

Yang, K. (2010). *Making Sense of Statistical Methods in Social Research*. Los Angeles: SAG.

Ye, X., Xiao, X., Shi, J. & Ling, M. (2016). The New Concepts of Measurement Error Theory. *Measurement, 83*, 96-105. DOI: 10.1016/j.measurement.2016.01.038

Youssef, M. A. F. M. (2011). Effective sample size calculation: How many patients will I need to include in my study? *Middle East Fertility Society Journal,* 16(4), 295-296. DOI: 10.1016/j.mefs.2011.10.

ประวัติผู้วิจัย

นายเรวัต แสงสุริยงค์ อาจารย์ประจำภาควิชาสังคมวิทยา ตำแหน่งผู้ช่วยศาสตราจารย์ การศึกษาระดับปริญญาตรี การศึกษาบัณฑิต (สังคมศึกษา) จากมหาวิทยาลัยศรีนครินทรวิโรฒ บางแสน และศิลปศาสตรบัณฑิต (รัฐศาสตร์) จากมหาวิทยาลัยรามคำแหง ระดับปริญญาโท สังคมวิทยามหาบัณฑิต (สังคมวิทยา) จากจุฬาลงกรณ์มหาวิทยาลัย และระดับปริญญาเอก รัฐศาสตรดุษฎีบัณฑิต (รัฐศาสตร์) จากจุฬาลงกรณ์มหาวิทยาลัย

บรรณานุกรม

บรรณานุกรม

Call number	Degree name	Level	Discipline	Hypothesis	Population type	Sample size	Sampling type	Sampling method	Variable level	Basic analysis	Missing detected	Data transformations	Statistic type	Statistics

คณะกรรมการพิจารณาจริยธรรมการวิจัยในมนุษย์
มหาวิทยาลัยบูรพา

วันที่รับรอง : 10 ม.ค. 2561

ภาคผนวก ข

ผลการวิเคราะห์ความเที่ยงตรงเชิงโครงสร้างของแบบวัดความเสี่ยงของการเกิดความคลาดเคลื่อนในการวิจัยเชิงปริมาณด้านสังคมวิทยา

ที่ ๑/๒๕๖๑

เอกสารรับรองผลการพิจารณาจริยธรรมการวิจัยในมนุษย์
มหาวิทยาลัยบูรพา

คณะกรรมการพิจารณาจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยบูรพา ได้พิจารณาโครงการวิจัย

รหัสโครงการวิจัย Hu 001/2561
โครงการวิจัยเรื่อง ความเสี่ยงของการเกิดความคาดเคลื่อนในการวิจัยเชิงปริมาณด้านสังคมวิทยา
หัวหน้าโครงการวิจัย ผู้ช่วยศาสตราจารย์ ดร.เรวัต แสงสุริยงค์
หน่วยงานที่สังกัด คณะมนุษยศาสตร์และสังคมศาสตร์

คณะกรรมการพิจารณาจริยธรรมการวิจัยในมนุษย์ มหาวิทยาลัยบูรพา ได้พิจารณาแล้วเห็นว่าโครงการวิจัยดังกล่าวเป็นไปตามหลักการของจริยธรรมการวิจัยในมนุษย์ โดยที่ผู้วิจัยเคารพสิทธิและศักดิ์ศรีในความเป็นมนุษย์ ไม่มีการล่วงละเมิดสิทธิ สวัสดิภาพ และไม่ก่อให้เกิดภยันตรายแก่ตัวอย่างการวิจัยและผู้เข้าร่วมโครงการวิจัย

จึงเห็นสมควรให้ดำเนินการวิจัยในขอบข่ายของโครงการวิจัยที่เสนอได้ (ดูตามเอกสารตรวจสอบ)

๑. เอกสารโครงการวิจัยฉบับภาษาไทย ฉบับที่ ๑ วันที่ ๑๐ เดือน มกราคม พ.ศ. ๒๕๖๑
๒. เอกสารชี้แจงผู้เข้าร่วมโครงการวิจัย ฉบับที่ - วันที่ - เดือน - พ.ศ. -
๓. เอกสารแบบแสดงความยินยอมของผู้เข้าร่วมโครงการวิจัย
ฉบับที่ - วันที่ - เดือน - พ.ศ. -
๔. เอกสารแสดงรายละเอียดเครื่องมือที่ใช้ในการวิจัยซึ่งผ่านการพิจารณาจากผู้ทรงคุณวุฒิแล้ว หรือชุดที่ใช้เก็บข้อมูลจริงจากผู้เข้าร่วมโครงการวิจัย ฉบับที่ ๑ วันที่ ๑๐ เดือน มกราคม พ.ศ. ๒๕๖๑

การรับรองผลการพิจารณาจริยธรรมการวิจัยในมนุษย์ฉบับนี้ มีผลถึงวันที่ ๙ เดือน มกราคม พ.ศ. ๒๕๖๒

ออกให้ ณ วันที่ ๑๐ เดือน มกราคม พ.ศ. ๒๕๖๑

ลงนาม
(ผู้ช่วยศาสตราจารย์ ดร.วิทวัส แจ้งเอี่ยม)
ประธานคณะกรรมการพิจารณาจริยธรรมการวิจัยในมนุษย์
มหาวิทยาลัยบูรพา