

บทที่ ๑

บทนำ

คำนำ



ความเป็นมาของปัญหา

ในปัจจุบันนี้ ตัวเลขสถิติต่าง ๆ เข้ามามีบทบาทในชีวิตประจำวันมาก สังเกตได้จากการทำงานที่มีผู้ใช้ข้อมูลสถิติเป็นพื้นฐานช่วยในการวางแผนโครงการต่าง ๆ ทั้งของเอกชนและระดับชาติ เนื่องจากคนทั่วไปยอมรับแล้วว่า การรู้ถึงสถิติที่เกี่ยวข้องของอนาคตข้างหน้า และสถิติข้อเท็จจริงในปัจจุบันจะเป็นประโยชน์ในการตัดสินใจทำแผนงานต่าง ๆ ให้ได้ผลดีที่สุด และลดโอกาสผิดพลาดใ้มากที่สุด โดยเฉพาะข้อมูลที่ได้จากการสำมะโนประชากรและเคหะ มีรายละเอียดที่สมบูรณ์ ถูกต้อง และเชื่อถือได้มากที่สุด จึงได้มีผู้มาขอทราบสถิติต่าง ๆ จากทางสำนักงานสถิติแห่งชาติมากมาย อย่างไรก็ตาม เมื่อผู้ใช้ใดข้อมูลที่ต้องการแล้วก็มักจะมีความสนใจที่จะถามเกี่ยวกับอัตราความถูกต้องของข้อมูล ว่าข้อมูลที่ได้อาจผิดพลาดไปจากความจริงเท่าใด และข้อมูลจากการสำมะโนมีความเชื่อถือได้เพียงใด ซึ่งมักจะแสดงเป็นอัตรารอยละของความคลาดเคลื่อนของข้อมูล และอาจวัดความแปรปรวน (variance) และความเอนเอียง (bias) ของข้อมูล

นอกจากนี้ในการเริ่มวางแผนสำมะโนประชากรแต่ละครั้งจะต้องอาศัยรายงานการทำสำมะโนครั้งก่อน ๆ และรายงานการสำรวจภายหลังการแจงนับสำมะโน P.E.S. (Post Enumeration Survey) ซึ่งจะแสดงถึงการวิเคราะห์ความบกพร่องของการทำสำมะโน ปกติจะระบุขนาดของความคลาดเคลื่อน (error) จากความจริง และปรับปรุง (adjust) ข้อมูลจากการสำมะโนไว้ด้วย เพื่อจะได้มีโอกาสทราบและแก้ไขข้อบกพร่องต่าง ๆ และปัญหาต่าง ๆ ที่เกิดขึ้น เพื่อนำเอาผลที่ได้เป็นแนวทางปรับปรุงให้การทำสำมะโนครั้งต่อไปสมบูรณ์ยิ่งขึ้น เนื่องจากเท่าที่ทราบยังไม่เคยมีรายงานวิเคราะห์ความคลาดเคลื่อนของการสำมะโนเลย จึงเห็นควรที่จะทำการศึกษาดังเทคนิคต่าง ๆ ของสถิติที่ใช้ในการประเมินความคลาดเคลื่อนของสำมะโนขนาดใหญ่ และจะใช้ข้อมูลจากการสำมะโนประชากรและเคหะ

พ.ศ. ๒๕๑๓ มาวิเคราะห์หาความคลาดเคลื่อนทำเป็นรายงานไว้ เพื่อเป็นประโยชน์แก่ผู้ใช้ ข้อมูลสำมะโนประชากรให้ใคร่และระมัดระวังถึงความคลาดเคลื่อนของสถิตินั้นตามสมควร และอาจเป็นประโยชน์ต่อการทำสำมะโนครั้งต่อไปอีกด้วย

ในการทำสำมะโนประชากรและเคหะ พ.ศ. ๒๕๑๓ เป็นครั้งแรกที่ได้มีการวางแผน และทำการสำรวจภายหลังการแจงนับ (P.E.S.) ทั้งด้านทฤษฎีและงานสนาม ซึ่งทางหน่วย วิเคราะห์ข้อมูล กองวิชาการสถิติ สำนักงานสถิติแห่งชาติ ได้ดำเนินการรวบรวมผลความแตกต่างระหว่างการทำสำมะโนและ P.E.S. แล้ว แต่ยังมีใ้มีการวิเคราะห์ถึงผลความคลาดเคลื่อนอย่างละเอียด จึงได้ขอคำแนะนำถึงวิธีการวิเคราะห์ และติดต่อขอข้อมูลมาวิเคราะห์ หาความคลาดเคลื่อนที่จะเป็นประโยชน์แก่ผู้ที่เกี่ยวข้อง ทำให้การวิเคราะห์สำเร็จลงและได้รับผลเป็นที่น่าพอใจ

วัตถุประสงค์และขอบเขตของการวิจัย

๑. ศึกษาค้นคว้าเกี่ยวกับระเบียบวิธีสถิติที่จะใช้ในการวัดและประมาณขนาดและประเภทของความคลาดเคลื่อน (error) ของข้อมูลที่ไ้จากการทำสำมะโนและสำรวจขนาดใหญ่ และระเบียบวิธีสถิติที่ใช้ในการวัดความคลาดเคลื่อนที่ใช้ในประเทศที่พัฒนาแล้ว เช่น ประเทศสหรัฐอเมริกา แคนาดา และประเทศที่กำลังพัฒนา เช่น ประเทศในอาฟริกา เพื่อจะได้ทราบถึงปัญหาและเหตุผลที่ใช้วิธีการวัดต่างกันในแต่ละประเทศ เพื่อจะหาวิธีการที่เหมาะสมมาใช้สำหรับประเทศไทย
๒. นำผลความคลาดเคลื่อนที่ประเมินได้มาใช้ในการประเมินการสำมะโนประชากรและเคหะ พ.ศ. ๒๕๑๓ ของประเทศไทย เพื่อตรวจสอบความครบถ้วนของการนับจดและแจงนับของการทำสำมะโนว่ามีความคลาดเคลื่อนจากข้อเท็จจริงเพียงใด และคำนวณหาความคลาดเคลื่อนของคู่รวมในการสำมะโนและความคลาดเคลื่อนของรายละเอียดที่สำคัญบางรายการ เป็นการประเมินขนาดและชนิดของความคลาดเคลื่อน ตลอดจนมูลเหตุที่ทำให้เกิดความคลาดเคลื่อนนั้น ๆ ขึ้น

๓. ตรวจสอบความเชื่อถือได้และความเหมาะสมของระเบียบวิธีนั้น เพื่อจะสามารถนำผลที่วิเคราะห์ได้ไปใช้ประโยชน์ และอาจเป็นแนวทางช่วยปรับปรุงผลการสำมะโนประชากรให้ดีขึ้น จะได้นั่นใจได้ว่าวิธีการวิเคราะห์ที่ใช้นั้นถูกต้อง

ประโยชน์ที่จะได้รับจากการวิจัย

๑. ทำให้ได้ระเบียบวิธีที่เหมาะสมสำหรับวัดและประมาณขนาดและประเภทความคลาดเคลื่อน อันจะเป็นประโยชน์แก่ผู้ที่ศึกษาวิเคราะห์เรื่องความคลาดเคลื่อนของข้อมูลอื่น ๆ ต่อไป เนื่องจากยังมิได้มีการศึกษาค้นคว้าเรื่องนี้เท่าที่ควร

๒. ทำให้ทราบความเชื่อถือได้และความเหมาะสมของระเบียบวิธีการนี้ ซึ่งผู้สนใจอาจค้นคว้าเพิ่มเติมเพื่อหาระเบียบวิธีการวัดความคลาดเคลื่อนที่อาจมีความเหมาะสมหรือเชื่อถือได้ยิ่งขึ้น อันจะเป็นประโยชน์แก่วงการสถิติของประเทศต่อไป

๓. จากรายงานผลความคลาดเคลื่อนต่าง ๆ ที่วัดได้ จะทำให้ผู้ใช้สถิติสำมะโนได้ทราบว่าการทำสำมะโนประชากรและเคหะ พ.ศ. ๒๕๑๓ นี้มีความคลาดเคลื่อนอะไรเกิดขึ้นบ้าง มากน้อยเพียงใด จะสามารถปรับปรุงผลการสำมะโนให้ถูกต้องได้อย่างไร และการนำผลการวัดความเที่ยงตรงของตัวเลขสำมะโนไปประกอบกับผลการสำมะโนจะเป็นแนวทางช่วยในการพิจารณาใช้ข้อมูลสำมะโนในทางที่สมควร (valid way) เป็นเครื่องเตือนใจไม่ให้มีการใช้ข้อมูลสำมะโนไปช่วยกำหนดนโยบายใด ๆ โดยเชื่อว่าข้อมูลนั้นเที่ยงตรง (accurate) กว่าที่เป็นจริง

๔. ผลที่ได้จะใช้เป็นแนวทางในการปรับปรุงการสำมะโนประชากรครั้งต่อไปให้สมบูรณ์ยิ่งขึ้น การวัดชนิดและอัตราความคลาดเคลื่อนจะให้ข้อมูลเพิ่มเติมที่จะใช้ปรับปรุงการทำสำมะโนครั้งต่อไป จากการระบุนขนาดและรู้ถึงเหตุผลของความคลาดเคลื่อนแต่ละชนิดจะทำให้คัดลึนใจได้ว่าจะใช้วิธีการวัด (measure) อะไรจึงจะช่วยลดหรือทำให้ความคลาดเคลื่อนนั้น ๆ หมดไป

๕. มีความสำคัญมากต่อการนำข้อมูลไปใช้ในการวางแผนพัฒนาเศรษฐกิจและสังคมแห่งชาติ ซึ่งอาศัยข้อมูลจากการสำมะโนประชากรเป็นพื้นฐานสำหรับการวางแผนต่าง ๆ โดยเฉพาะข้อมูลแสดงรายละเอียดของลักษณะต่าง ๆ ของประชากรในประเทศ



วิธีการดำเนินการคนควาและวิจัย

๑. ศึกษาทฤษฎีและระเบียบวิธีที่ต่างกันในการประเมินความคลาดเคลื่อนการทำสำมะโนขนาดใหญ่ที่ไ้ทำกันในประเทศต่าง ๆ รวมทั้งในประเทศไทย

๒. ศึกษาวิเคราะห์แผนงานสำมะโนประชากรและเคหะของประเทศไทย พ.ศ. ๒๕๑๓ รวมทั้งวิธีการรวบรวมและวิเคราะห์ผลการสำมะโนประชากรของสำนักงานสถิติแห่งชาติ

๓. นำทฤษฎีที่ศึกษาไว้มานำประยุกต์กับข้อมูลการสำมะโน ศึกษาแผนงานการสำรวจภายหลังการแจงนับสำมะโนประชากรและเคหะ พ.ศ. ๒๕๑๓ (P.E.S.) นำข้อมูลที่ทางหน่วยวิเคราะห์ข้อมูล กองวิชาการสถิติ สำนักงานสถิติแห่งชาติรวบรวมไว้มานำวิเคราะห์เปรียบเทียบกับข้อมูลสำมะโน โดยใช้อข้อมูล P.E.S. เป็นหลักในการสอบเทียบ (matching) หาความแตกต่างระหว่างการสำรวจทั้ง ๒ ครั้ง และนำผลที่ได้ไปวิเคราะห์ตามทฤษฎีที่กล่าวแล้ว เพื่อประเมินค่าความคลาดเคลื่อนของการทำสำมะโนและประเมินผลในบางรายการที่สำคัญ ๆ นอกจากนี้จะได้เสนอแนะถึงวิธีการที่อาจใช้ปรับปรุงตัวเลขสำมะโน และการลดความคลาดเคลื่อนอีกด้วย

๔. สรุปผลสืบเนื่องของระเบียบวิธีที่ใช้

องค์ประกอบที่สำคัญของความคลาดเคลื่อนของข้อมูลสถิติ
ที่ไ้จากการทำสำมะโนหรือการสำรวจ

ในการทำสำมะโนประชากรหรือการสำรวจต่าง ๆ ผลของการสำรวจในทางสถิติถือว่าอาจมีความคลาดเคลื่อนจากความเป็นจริงไปไ้ด้วยสาเหตุหลายประการ ความคลาดเคลื่อนในการทำสำมะโนแยกออกเป็น

๑. ความคลาดเคลื่อนจากงานสนาม (field work) เกิดขึ้นจากวิธีการเก็บ

และรวบรวม และประมวลผลข้อมูล เช่น

๑.๑ การนับจนบุคคลลงในชั้นที่ไม่ถูกต้อง (misclassifying people)

ในระหว่างการสัมภาษณ์

๑.๒ ความคลาดเคลื่อนที่เกิดในระหว่างการบันทึกข้อมูลสำมะโนจากรายการสำมะโนซึ่งกรอกโดยผู้ตอบสัมภาษณ์ลงในแบบบันทึกข้อมูล

๑.๓ ขั้นตอนประมวลผลก็อาจมีความคลาดเคลื่อนในระหว่างการบรรณาธิกรณ (edit) ทั้งด้วยมือและเครื่องจักร ซึ่งจะต้องผ่านขั้นตอนการรวบรวมและลำดับเรื่องใหม่ และในการให้รหัส (coding) และการประมวลผล (processing) ก็อาจเกิดความคลาดเคลื่อนขึ้นได้

๒. ความคลาดเคลื่อนที่เกิดจากข้อมูลเอง ข้อมูลที่ได้รับนั้นยังมีความคลาดเคลื่อนที่เกิดจาก

๒.๑ ความคลาดเคลื่อนจากการสุ่มตัวอย่าง (sampling error)

๒.๒ ความคลาดเคลื่อนอื่น ๆ ที่ไม่ได้เนื่องจากการสุ่มตัวอย่าง (non-sampling error)

การหาความคลาดเคลื่อนทางสถิติในวิทยานิพนธ์ฉบับนี้จะมุ่งศึกษาเฉพาะความคลาดเคลื่อนอื่นเนื่องมาจากสาเหตุอื่นที่ไม่ใช่จากการสุ่มตัวอย่าง (non-sampling error) ส่วนความคลาดเคลื่อนจากการสุ่มตัวอย่างจะไม่นำมาพิจารณา เนื่องจากการทำสำมะโนประชากรหรืองานสำรวจขนาดใหญ่มักจะสำรวจประชากรทั้งหมด ถึงแม้จะมีการสุ่มตัวอย่างก็เป็นตัวอย่างขนาดใหญ่จนความคลาดเคลื่อนจากการสุ่มตัวอย่างไม่มี หรือมีก็น้อยมากจนอาจตัดทิ้ง (eliminate) ได้

ความคลาดเคลื่อนของข้อมูลสถิติที่มิได้เกิดจากการสุ่มตัวอย่าง (non-sampling error) ของสำมะโนขนาดใหญ่

ความคลาดเคลื่อนนี้จะพยายามวัดประกอบด้วย

ความคลาดเคลื่อนของคัมรวม (coverage error)

เป็นความคลาดเคลื่อนในยอดรวมของประชากรหรือเคหะ เกิดขึ้นจากการที่จำนวน

๑ คัมรวม (coverage) หมายถึงขอบข่ายของการทำสำมะโนซึ่งได้กำหนดประเภทของประชากรที่จะต้องแจงนับ

ครัวเรือนหรือจำนวนคนที่นับมาได้ขาดหายไป หรือนับเกินมาจากความเป็นจริง ในขณะที่ทำการแจงนับในสนาม เป็นผลเนื่องจากการที่ไม่อาจทำการสัมภาษณ์ทุกหน่วย (unit) ในหน่วยประชากร (population unit) ละ ๑ ครั้งเท่านั้น ทั้งนี้เพราะการสำมะโนเป็นงานระดับชาติ ต้องอาศัยความร่วมมือจากพนักงานแจงนับ (interviewer) จำนวนมาก จึงมีการกำหนดขอบเขตการทำสำมะโนของแต่ละคนเป็นเขตแจงนับ ซึ่งเรียกว่า E.D. (enumeration district) แบ่งตามพื้นที่ จึงเกิดปัญหาเรื่องความไม่แน่ชัดของเส้นแบ่งเขตแจงนับ ความคลาดเคลื่อนที่เกิดขึ้นแบ่งออกเป็น

๑. แจงนับขาด (non coverage) หรือการที่ประชากรหรือครัวเรือนขาดหายไป (omitted) ความคลาดเคลื่อนนี้สามารถวัดได้โดยการเลือกตัวอย่างของเขตแจงนับ (E.D.) และใช้ความพยายามในการแจงนับรายชื่อประชากรและครัวเรือนทั้งหมดในเขตนั้นซ้ำอีกทีหนึ่งในช่วงเวลาที่มีการสำรวจสำมะโน โดยใช้วิธีสอบเทียบรายชื่อหนึ่งจากการแจงนับทั้งสองครั้งที่ละรายชื่อ (one-to-one matching) จนครบทุกรายชื่อเพื่อหาความแตกต่างระหว่างการแจงนับทั้งสองครั้ง เมื่อมีความแตกต่างอาจต้องทำงานสนาม (field work) คือกลับไปสัมภาษณ์หน่วยที่สงสัยนั้นอีกทีเพื่อจะตกลงว่าคำตอบใดถูกต้อง และจะสามารถไขข้อสงสัยถึงเหตุผลสำหรับความคลาดเคลื่อนที่เกิดขึ้นนั้น

๒. แจงนับเกิน (erroneous inclusion) หรือการที่มีประชากรหรือครัวเรือนเพิ่มขึ้นจากความเป็นจริง เป็นการแจงนับหน่วยใดเกิน ๑ ครั้ง (replication or duplicate) ซึ่งอาจเกิดจากความบังเอิญนับซ้ำ เพราะเหตุที่แผนที่ไม่ถูกต้องสมบูรณ์พอ ทำให้มีเขตแจงนับบางอันเหลื่อมกัน (overlapping enumeration areas) หรือเกิดจากการตีความในคำจำกัดความในสำมะโนผิดพลาดไป (misinterpret of census definition) ทำให้แจงนับบุคคลที่ไม่เข้าชายที่จะแจงนับ หรือเกิดจากจำนวนเด็กซึ่งเกิดหลังจากวันสำมะโนแต่เสียชีวิตก่อนวันที่ถูกสัมภาษณ์ได้ถูกนับรวมเข้ามาในจำนวนบุตร เกิดรอยคดควาย

ความคลาดเคลื่อนในเนื้อหา (Content error)

คือความคลาดเคลื่อนในข้อเท็จจริงหรือเนื้อหาที่ถูกรายงานสำหรับหน่วยที่ถูกแจงนับ

โดยเน้นถึงความคลาดเคลื่อนของคำตอบ (response error) ซึ่งเกิดขึ้นเนื่องจากการที่พนักงานแจกแบบสัมภาษณ์ประชากรเป็นรายบุคคลแล้วเกิดนึกพลาดไป คือผู้ตอบอาจจะตอบผิดหรือการบันทึกผิดพลาดได้ บางครั้งเป็นเพราะผู้ตอบไม่เข้าใจคำถาม หรืออาจไม่อยากจะตอบในบางเรื่อง จึงอาจลืมตอบคำถามบางข้อหรือให้คำตอบที่ไม่สมบูรณ์ ทำให้ผลที่ได้คลาดเคลื่อนจากความเป็นจริง ความคลาดเคลื่อนของคำตอบสามารถวัดได้ในเทอมของ

๑. ความเอนเอียงของคำตอบของข้อมูลสถิติ (response bias of the statistics) เป็นผลต่างระหว่างค่าเฉลี่ยซึ่งได้มาจากการทำการสำรวจซ้ำหลาย ๆ ครั้ง กับค่าเฉลี่ยซึ่งจะถูกวัดจากการสำรวจซึ่งดำเนินการโดยไม่มี ความคลาดเคลื่อนซึ่งมักใช้แทนด้วยค่าเฉลี่ยจาก P.E.S.

๒. ความแปรปรวนของคำตอบของข้อมูลสถิติซึ่งเกิดจากการสำรวจอันใดอันหนึ่ง เป็นความแปรปรวนระหว่างค่าของการกระทำครั้งหนึ่ง (trial) กับการกระทำหรือการสำรวจเป็นจำนวนหลาย ๆ ครั้ง

การสำมะโนขนาดใหญ่

ในประเทศไทยได้มีการจัดทำสำมะโนขนาดใหญ่ขึ้นหลายครั้ง เพื่อจะศึกษาหาข้อมูลเกี่ยวกับคุณลักษณะ โครงสร้าง และความต้องการของประชากรทั่วประเทศ อันจะเป็นประโยชน์ในการวางแผนพัฒนาประเทศ ทั้งทางด้านเศรษฐกิจและสังคม การสำรวจและสำมะโนที่ทำกันบ่อย ๆ ได้แก่

การสำรวจผลผลิตการเกษตรของประเทศ (Agricultural Production Survey)

การสำรวจภาวะการใช้แรงงานของประชากร (Labor Force Survey)

การสำมะโนธุรกิจการค้าหรือบริการ (Census of Business Trades or Services)

การสำมะโนประชากรและเคหะ (Census of Population and Housing)

การสำรวจใหญ่ ๆ จะทำโดยสำนักงานสถิติแห่งชาติ ซึ่งมีหน้าที่รับผิดชอบในการเก็บและรวบรวมสถิติต่าง ๆ ของประเทศไทย ส่วนการสำรวจเล็ก ๆ หรือการวิจัยซึ่งใช้ตัวอย่าง

ไม่มากนักอาจทำขึ้นจากหน่วยงานที่กองการข้อมูลเหล่านั้นโดยตรงก็มีบ้าง การสำมะโนที่สำคัญและได้มาตรฐานที่สุดก็คือการสำมะโนประชากรและเคหะ ดังนั้นในวิทยานิพนธ์ฉบับนี้จะเป็นการศึกษาวิเคราะห์โดยอาศัยข้อมูลจากการสำมะโนประชากรและเคหะ พ.ศ. ๒๕๑๓ เป็นสำคัญ

การสำมะโนประชากรและเคหะ^๒ คือการรวบรวมข้อมูลเกี่ยวกับบุคคลทุกคนซึ่งอยู่รวมกันเป็นหน่วยเรียกว่าครัวเรือน และเกี่ยวกับที่อยู่อาศัยของครัวเรือนเหล่านั้น ข้อมูลที่ได้นี้จะเป็นข้อเท็จจริงที่เป็นอยู่ในวันสำมะโน (census date) ที่กำหนดไว้เท่านั้น ฉะนั้นการทำสำมะโนประชากรและเคหะจึงเป็นเสมือนการถ่ายภาพนิ่งของประชากรและเคหะของประเทศในขณะเวลาหนึ่ง ซึ่งแสดงให้เห็นจำนวนและลักษณะต่าง ๆ ของประชากรและเคหะตามที่ เป็นจริงในขณะนั้น

การสำมะโนประชากรและเคหะเป็นงานที่สำคัญมาก เนื่องจากทำการแจกนับประชากรทุกคนทั่วราชอาณาจักร ทำให้ทราบถึงรายละเอียดที่สำคัญ ๆ เช่น จำนวนประชากรและจำนวนครัวเรือนในท้องที่ต่าง ๆ ทั่วประเทศ แสดงถึงเพศ อายุ การศึกษา สถานภาพการทำงาน อาชีพ และข้อมูลเกี่ยวกับที่อยู่อาศัยของประชากร ซึ่งจะได้นำข้อมูลเหล่านั้นมาใช้ประกอบการพิจารณาวางโครงการต่าง ๆ เป็นประโยชน์ต่อเอกชนในการวางแผนการผลิต การตลาด และเป็นประโยชน์แก่ทางราชการที่จะวางแผนพัฒนาเศรษฐกิจและสังคมเพื่อความกินดีอยู่ดีของประชาชน จึงควรให้ความสนใจในการวิเคราะห์หาความคลาดเคลื่อนของการสำมะโนประชากรและเคหะ เพื่อใช้ประกอบกับผลการสำมะโน ทำให้ได้ประโยชน์มากขึ้นไปกว่าเดิม

^๒ สำนักงานสถิติแห่งชาติ, แผนงานจัดทำสำมะโนประชากรและเคหะ พ.ศ. ๒๕๑๓ ของประเทศไทย (๒๕๑๓), หน้า ๑