สำนักงานสถิติแห่งชาติ

(ร่าง) คู่มือการพยากรณ์ย้อนหลัง

BACKCASTING HANDBOOK

Based on Implementation Guide
for the International Standard Industrial Classification (ISIC), Rev.4
United Nations Statistical Division

หน่วยงานเจ้าของเรื่อง กองบริหารจัดการระบบสถิติ
สำนักงานสถิติแห่งชาติ
โทรศัพท์ 0 2141 7447-9
โทรสาร 0 2143 8121
ไปรษณีย์อิเล็กทรอนิกส์ : pstand@nso.go.th

หน่วยงานที่เผยแพร่ กองสถิติพยากรณ์
สำนักงานสถิติแห่งชาติ
ศูนย์ราชการเฉลิมพระเกียรติฯ 80 พรรษา อาคารรัฐประศาสนภักดี ชั้น 2
ถนนแจ้งวัฒนะ เขตหลักสี่ กรุงเทพฯ 10210
โทรศัพท์ 0 2141 7498
โทรสาร 0 2143 8132
ไปรษณีอิเล็กทรอนิกส์ binfopub@nso.go.th

ปีที่จัดพิมพ์

คำนำ

การเปลี่ยนการจัดจำแนก และการเปลี่ยนแปลงคำนิยามของข้อมูล รวมถึงการเปลี่ยนแปลงวิธีการต่าง ๆ ในระเบียบวิธี ทำให้อนุกรมข้อมูลเกิดความเสียหาย ส่งผลให้ข้อมูลสองช่วงเวลาทั้งก่อนและหลังการเปลี่ยนแปลงไม่สามารถนำมาเปรียบเทียบหรือนำมาวิเคราะห์ร่วมกันได้ รวมถึงส่งผลกระทบต่อผู้ใช้งานที่ต้องใช้อนุกรมระยะยาวสำหรับการวิเคราะห์ หรือคาดการณ์ข้อมูล

สำนักงานสถิติแห่งชาติในฐานะที่เป็นหน่วยงานที่มีหน้าที่ผลิตสถิติและบริหารจัดการระบบสถิติ รวมถึงพัฒนามาตรฐานสถิติ ตามแผนแม่บทระบบสถิติประเทศไทย ฉบับที่ 2 (พ.ศ. 2559-2564) ที่มุ่งเน้นในการพัฒนาคุณภาพสถิติให้ได้มาตรฐานสากล ได้ตระหนักถึงปัญหาที่เกิดจากการเปลี่ยนแปลงการจัดจำแนกดังกล่าว ดังนั้น จึงได้จัดทำคู่มือการพยากรณ์ย้อนหลังข้อมูล (Backcasting) ขึ้น เพื่อให้ผู้ผลิตข้อมูลใช้เป็นแนวทางในการจัดการข้อมูล เมื่อมีการปรับปรุงการจัดจำแนกหรือปรับปรุงคำนิยามของมูล ด้วยวิธีการพยากรณ์ย้อนหลังให้ข้อมูลก่อนการเปลี่ยนแปลงให้อยู่ในรูปของคำนิยามหรือการจัดจำแนกใหม่ เพื่อลดผลกระทบจากการแบ่งช่วงของอนุกรมลงและสามารถเปรียบเทียบข้ามช่วงเวลาได้

คู่มือฉบับนี้ได้อ้างอิงวิธีการพยากรณ์ย้อนหลังจากเอกสาร Implementation Guide for the International Standard Industrial Classification (ISIC), Rev.4 ที่จัดทำโดยองค์การสถิติแห่งสหประชาชาติ (United Nations Statistics Division: UNSD) โดยรวบรวมตั้งแต่การวางแผนในการนำ ISIC Rev.4 มาใช้ การปรับใช้ ISIC Rev.4 ให้เหมาะสมสำหรับแต่ละประเทศ วิธีการในการออกแบบการสุ่มตัวอย่างและการประมาณค่าถ่วงน้ำหนัก รวมถึงการพยากรณ์ย้อนหลังข้อมูล สำหรับคู่มือการพยากรณ์ย้อนหลังฉบับนี้ ได้นำเอกสารดังกล่าวมาแปลเป็นภาษาไทยเฉพาะส่วนที่เกี่ยวข้องกับการพยากรณ์ย้อนหลังเท่านั้น เพื่อเป็นแนวทางให้ผู้ผลิตข้อมูลนำวิธีการพยากรณ์ย้อนหลังไปปรับใช้ได้

สารบัญ

หน้า

คำนำ **ค**

สารบัญ **จ**

สารบัญตาราง **ฉ**

การพยากรณ์ย้อนหลัง **1**

บทนำ **3**

วิธีการพยากรณ์ย้อนหลัง **7**

1. วิธีการที่อ้างอิงตามการปรับปรุงข้อมูลรายบุคคล (วิธีข้อมูลระดับจุลภาค) 9
   - 1.1 วิธีข้อมูลระดับจุลภาคคืออะไร 9
   - 1.2 ประเด็นการจัดจำแนกที่เกี่ยวกับวิธีข้อมูลระดับจุลภาค 10
2. วิธีที่อ้างอิงตามค่าสัมประสิทธิ์การเปลี่ยนแปลง (วิธีสัดส่วน) 17
   - 2.1 วิธีสัดส่วนคืออะไร 17
   - 2.2 การใช้วิธีสัดส่วน 19
3. ข้อดีและข้อเสียของวิธีข้อมูลระดับจุลภาคและวิธีสัดส่วน 23
4. การใช้วิธีประมาณค่าในช่วงระหว่างปีมาตรฐาน (การรวมวิธีข้อมูลระดับจุลภาคและวิธีสัดส่วน) 27
5. ตัวอย่าง 29

สารบัญตาราง

หน้า

ตาราง 1	วิธีการคำนวณการจัดสรรปัจจัยสำหรับ SYPRO	30
ตาราง 2	การสร้างปัจจัยการเปลี่ยนแปลงของ SYPRO	31
ตาราง 3	ตัวอย่างการคำนวณปัจจัยการเปลี่ยนแปลง	31
ตาราง 4	การนำปัจจัยการเปลี่ยนแปลงไปใช้ (1) ปรับฐานดัชนี SYPRO	32
ตาราง 5	การนำปัจจัยการเปลี่ยนแปลงไปใช้ (2) รวมดัชนี SYPRO ที่ปรับฐานแล้ว	32

การพยากรณ์ย้อนหลัง
BACKCASTING

Based on Implementation Guide
for the International Standard Industrial
Classification (ISIC), Rev.4

United Nations Statistical Division

บทนำ

ความเป็นมาและขอบเขตของคู่มือ

ข้อมูลอนุกรมเวลา คือข้อมูลที่มีการบันทึกไว้ในช่วงเวลาที่ต่อเนื่องกัน ซึ่งช่วงเวลาอาจเป็นรายวัน รายสัปดาห์ รายเดือน รายไตรมาส หรือรายปี ข้อมูลอนุกรมเวลาอาจมีลักษณะขึ้น-ลง หรือราบเรียบ โดยการขึ้นหรือลงของข้อมูลเกิดจากความผันแปรของข้อมล เช่น ความผันแปรตามฤดูกาล การผันแปร ตามวัฏจักร เป็นต้น อนุกรมระยะยาวมีความสำคัญสำหรับผู้ใช้งานสถิติ เนื่องจากอนุกรมเวลาช่วยให้ผู้ใช้เห็นถึงการเปลี่ยนแปลงและแนวโน้มของเรื่องนั้น ๆ ช่วยกำหนดรูปแบบการผันแปร และยังให้ข้อมูลที่จำเป็นสำหรับนักวิเคราะห์ในการพยากรณ์หรือคาดการณ์ผลลัพธ์ในอนาคต รวมถึงยังช่วยให้ผู้บริหารและผู้กำหนดนโยบายสามารถใช้ในการคาดการณ์และวางแผนในอนาคตได้ อย่างไรก็ตาม ในการรวบรวมข้อมูลในบางครั้งจำเป็นต้องมีการปรับปรุงคำนิยามของข้อมูล การจัดจำแนกของข้อมูล รวมถึงการเปลี่ยนแปลงในระเบียบวิธีต่าง ๆ เพื่อรองรับและตอบสนองกับการเปลี่ยนแปลงที่เกิดขึ้นในปัจจุบัน แต่เมื่อมีการปรับปรุงหรือเปลี่ยนแปลงใด ๆ ของข้อมูล ทำให้ข้อมูลก่อนการเปลี่ยนแปลงและหลังการเปลี่ยนแปลงไม่สามารถนำมาใช้ร่วมกัน หรือใช้เปรียบเทียบกันได้ อีกทั้งยังส่งผลกระทบต่อข้อมูลอนุกรมที่ก่อให้เกิดการแบ่งช่วงของอนุกรม ทำให้ผู้ใช้งานอาจได้ผลการวิเคราะห์หรือการคาดการณ์คลาดเคลื่อน หากเกิดการเปลี่ยนแปลงข้อมูลในระดับประเทศที่สำคัญ เช่น ข้อมูลอุตสาหกรรม ข้อมูลแรงงาน ฯลฯ อาจส่งผลต่อการตัดสินใจของผู้บริหารหรือผู้กำหนดนโยบายที่ผิดพลาดได้ จากปัญหาดังกล่าว จึงควรแก้ไขโดยการปรับโครงสร้างของอนุกรมเวลาที่เคยใช้มาตรฐานการจัดจำแนกหรือนิยามที่ใช้อยู่เดิมให้เป็นไปตามมาตรฐานการจัดจำแนก หรือนิยามที่ปรับปรุงใหม่ ซึ่งวิธีการนี้เรียกว่า “การพยากรณ์ย้อนหลัง” (Backcasting)

วิธีการพยากรณ์ย้อนหลังเป็นวิธีการที่ตรงกันข้ามกับการพยากรณ์ข้อมูล (Forecasting) โดยการพยากรณ์ข้อมูลปกติจะทำเพื่อคาดคะเนข้อมูลในอนาคต โดยนำข้อมูลในอดีตหรือปัจจุบัน ณ ช่วงเวลาใดเวลาหนึ่งมาหาความสัมพันธ์ และหาแนวโน้มเพื่อคาดคะเนข้อมูลในอนาคต แต่การพยากรณ์ย้อนหลังมีวัตถุประสงค์เพื่อทำให้ข้อมูลที่แตกต่างกันสองช่วงเวลาสอดคล้องกัน โดยการใช้ความสัมพันธ์ของข้อมูลที่แตกต่างกันในปัจจุบัน ณ ช่วงเวลาใดเวลาหนึ่ง ทำการพยากรณ์ข้อมูลในอดีต

ในคู่มือนี้ได้รวบรวมวิธีการพยากรณ์ย้อนหลังมาจาก Implementation Guide for the International Standard Industrial Classification (ISIC), Rev.4 ที่จัดทำโดยองค์การสถิติแห่งสหประชาชาติ (United Nations Statistics Division: UNSD) เนื่องจากในปี ค.ศ. 2008

คณะกรรมาธิการด้านสถิติแห่งสหประชาชาติ (United Nations Statistical Commission: UNSC) ได้ประกาศให้มาตรฐานการจัดจำแนกอุตสาหกรรมตามกิจกรรมทางเศรษฐกิจในระดับสากล (International Standard Industrial Classification of All Economic Activities: ISIC) ฉบับปรับปรุงครั้งที่ 4 เป็นมาตรฐานสากลใหม่ที่ทุกประเทศสามารถนำไปปรับใช้แต่การที่แต่ละประเทศจะนำ NSIC [1] ที่ปรับปรุงมาใช้กับสถิติในระดับประเทศอาจส่งผลให้เกิดการแบ่งช่วงของอนุกรมเวลาที่เกิดจากการเปลี่ยนแปลงการจัดจำแนกตามที่กล่าวในข้างต้น เนื่องจากข้อมูลอุตสาหกรรมเป็นข้อมูลที่สำคัญในระดับประเทศ ใช้ในการคำนวณผลิตภัณฑ์มวลรวมของประเทศและใช้ในการวางแผน และกำหนดนโยบายที่สำคัญมากมาย เช่น การจัดทำบัญชีประชาชาติ ดังนั้น จึงได้จัดทำคู่มือการพยากรณ์ย้อนหลังขึ้น เพื่อให้แต่ละประเทศได้นำไปปรับใช้เมื่อมีการใช้ NSIC ฉบับปรุง รวมถึงสามารถนำไปใช้เมื่อมีการปรับปรุงการจัดจำแนกอื่น ๆ

เนื้อหาโดยสรุป

ในคู่มือการพยากรณ์ย้อนหลังนี้จะนำเสนอความรู้ วิธีการและเทคนิคในการพยากรณ์ย้อนหลังให้กับนักสถิติในการนำไปใช้ เมื่อมีการเปลี่ยนการจัดจำแนกในระบบสถิติ โดยได้นำเสนอวิธีการพยากรณ์ย้อนหลังวิธีต่าง ๆ พร้อมคำอธิบาย และตัวอย่าง รวมถึงข้อดีและข้อเสียของแต่ละวิธีการ โดยวิธีการการพยากรณ์ย้อนหลังที่จะนำเสนอดังต่อไปนี้ ไม่ได้มีเป้าหมายเพื่อกำหนดให้เป็นกฎเกณฑ์หรือวิธีการที่ทุกประเทศต้องปฏิบัติตาม เนื่องจากวิธีการเหล่านี้ไม่ได้มี “วิธีการที่ดีที่สุด” การเลือกใช้แต่ละวิธีขึ้นอยู่กับหลายปัจจัย ดังนี้

- ประเภทของอนุกรมสถิติที่ต้องการประมาณย้อนกลับ (ข้อมูลดิบ ข้อมูลรวมยอด (aggregate) ตัวชี้วัด อัตราการเจริญเติบโต ฯลฯ)
- สถิติหลักในอนุกรมเวลา
- ความพร้อมของข้อมูลระดับจุลภาค
- ความพร้อมของ “การลงรหัสคู่”(dual code) ในระดับข้อมูลจุลภาค (เช่น สถานประกอบการที่ถูกจัดจำแนกตามการจัดจำแนกเดิมและการจัดจำแนกใหม่)
- ระยะเวลาของ “การลงรหัสคู่”
- ความถี่ของอนุกรมเวลาที่มีอยู่
- รายละเอียดที่ต้องการประมาณย้อนกลับ และ
- การพิจารณาถึงค่าใช้จ่ายและทรัพยากร

[1] ในเอกสาร Implementation Guide for the International Standard Industrial Classification (ISIC), Rev.4 ได้มีการแนะนำว่าในการนำ ISIC Rev.4 ไปใช้ ควรมีการปรับรหัสให้เข้ากับบริบทของแต่ละประเทศ โดยควรแตกรหัสที่มีความสำคัญกับประเทศออกมาให้ละเอียดมากขึ้นและสอดคล้องกับความต้องการ แต่ยังคงสามารถเปรียบเทียบกับ ISIC ได้ โดยเรียก ISIC ที่ปรับในระดับประเทศว่า NSIC

ทุกวิธีการที่นำเสนอในคู่มือนี้ สมมติให้มีการลงรหัสคู่ (รหัสอุตสาหกรรมตามการจัดจำแนกเดิมและใหม่) ให้กับทุกหน่วยธุรกิจที่อยู่ในฐานข้อมูลการจดทะเบียนธุรกิจ (Business Register: BR) เป็นเวลาอย่างน้อยหนึ่งช่วงเวลา ซึ่งการกำหนดระยะเวลาในการลงรหัสคู่ขึ้นอยู่กับการตัดสินใจของแต่ละประเทศ

วิธีการพยากรณ์ย้อนหลังที่นำเสนอในหัวข้อที่ 1 เรียกว่า *"วิธีข้อมูลระดับจุลภาค"* โดยมีแนวคิดพื้นฐานคือ อนุกรมเวลาในอดีตจะถูกเปลี่ยนรหัส โดยรหัสที่เปลี่ยนอ้างอิงจากข้อมูลระดับจุลภาคใน BR หรือจากฐานข้อมูลรายบุคคลที่มีการลงรหัสคู่ในปัจจุบัน กล่าวคือ วิธีข้อมูลระดับจุลภาค คือการนำรหัสกิจกรรมใหม่มาแทนที่รหัสกิจกรรมเดิม และทำการคำนวณชุดอนุกรมเวลาในอดีตใหม่อีกครั้งตามรหัสที่มีการเปลี่ยนแปลง วิธีการนี้เป็นการปรับโครงสร้างอนุกรมเวลาที่มีความน่าเชื่อถือที่สุด แต่มีข้อเสียคือต้นทุนสูงและสัมประสิทธิ์การแปรผันสูง

วิธีการที่นำเสนอในหัวข้อที่ 2 คือ "วิธีสัดส่วน" ซึ่งเป็นการใช้ประโยชน์จาก "เมทริกซ์การเปลี่ยนแปลง" โดยเป็นวิธีการที่ยอมให้มีการเปลี่ยนแปลงข้อมูลในระดับรวมยอดที่จัดจำแนกตาม NSIC เดิม ให้อยู่ในรูปการจัดจำแนกตาม NSIC ใหม่ โดยที่เมทริกซ์การเปลี่ยนแปลง จะถูกคำนวณ ณ จุดหนึ่งของช่วงเวลาและใช้ตัวแปรอย่างน้อย 1 ตัวแปรในการอ้างอิง (เช่น จำนวนของสถานประกอบการที่จำแนกตามการจัดจำแนกเดิมและการจัดจำแนกใหม่) ข้อดีของวิธีนี้คือใช้ทรัพยากรและใช้เวลาน้อยกว่าวิธีข้อมูลระดับจุลภาค แต่มีข้อเสียคือความแม่นยำน้อยกว่าวิธีข้อมูลระดับจุลภาค

วิธีการที่นำเสนอในหัวข้อที่ 3 เป็นวิธีที่รวม "วิธีข้อมูลระดับจุลภาค" และ "วิธีสัดส่วน" เข้าด้วยกัน โดยต้องลงรหัสคู่ให้แต่ละหน่วยสถิติเป็นเวลามากกว่า 1 ปี และใช้วิธีประมาณค่าในช่วง (interpolating) ให้กับข้อมูลระหว่างช่วงเวลาที่ลงรหัสคู่ วิธีการนี้เป็นทางเลือกในการแก้ปัญหาวิธีที่นำเสนอในหัวข้อที่ 1 และ 2 ทั้งในด้านต้นทุนและคุณภาพของการปรับโครงสร้างอนุกรมเวลาใหม่

วิธีการพยากรณ์ย้อนหลัง

BACKCASTING METHODS

1 วิธีการที่อิงตามการปรับปรุงข้อมูลรายบุคคล (วิธีข้อมูลระดับจุลภาค)

1.1 วิธีข้อมูลระดับจุลภาคคืออะไร

เมื่อมีการเปลี่ยนแปลงการจัดจำแนกอุตสาหกรรม การใช้วิธีการ “ปรับปรุงข้อมูลระดับจุลภาค” คือการลงรหัสกิจกรรมใหม่ (ตามการจัดจำแนกใหม่) ให้กับแต่ละหน่วยสถิติในทุก ๆ ช่วงเวลาในอดีต จากนั้นจะทำการคำนวณข้อมูลรวมยอดให้แต่ละอนุกรมตามการจัดจำแนกใหม่

การใช้วิธีการนี้มีการดำเนินงานที่เฉพาะคือ จะดำเนินงานในระดับหน่วยสถิติ โดยทำการลงรหัสอุตสาหกรรมใหม่ให้สอดคล้องกับกิจกรรมหลัก (principle activity) และไม่ทำการเปลี่ยนแปลงข้อมูลรายบุคคลรายการอื่น ๆ ในฐานข้อมูล จากนั้นทำการรวมยอดข้อมูลอุตสาหกรรมโดยจัดกลุ่มอุตสาหกรรมต่าง ๆ ให้เป็นไปตามการจำแนกใหม่ ซึ่งคำนวณเช่นเดียวการคำนวณให้กับการจัดจำแนกเดิมในครั้งก่อน อย่างไรก็ตามขั้นตอนการรวมยอดข้อมูลจากการสำมะโนกับข้อมูลจากตัวอย่างนั้นมีความแตกต่างที่สำคัญในการใช้ค่าถ่วงน้ำหนัก ซึ่งควรแจ้งให้ผู้ใช้งานรับทราบ

ข้อแตกต่างระหว่างการใช้ค่าถ่วงน้ำหนักของข้อมูลจากการสำมะโนกับข้อมูลจากตัวอย่างคือ การสำมะโนหน่วยสถิติทั้งหมดจะถูกเลือกในการเก็บรวมรวมข้อมูล ดังนั้นจึงมีค่าถ่วงน้ำหนักเท่ากับหนึ่ง ดังนั้นในกรณีนี้ หลังจากเปลี่ยนรหัสใหม่ให้หน่วยสถิติ ในขั้นตอนการรวมยอดจึงสามารถทำได้โดยไม่ต้องคำนึงถึงค่าถ่วงน้ำหนัก แต่ในกรณีของการใช้ข้อมูลจากตัวอย่าง มีเพียงหน่วยตัวอย่างที่ถูกสุ่มเท่านั้นที่ถูกเปลี่ยนรหัส ดังนั้น ในการรวมยอดข้อมูลควรพิจารณาว่าจะประมาณค่าประชากรโดยใช้ค่าถ่วงน้ำหนักที่มาจากการสุ่มตัวอย่างในครั้งแรกหรือทำการปรับค่าถ่วงน้ำหนักใหม่ ซึ่งหากมีการปรับ หน่วยสถิติที่ไม่ได้ถูกสำรวจจะมีอิทธิพลจากการคำนวณค่าถ่วงน้ำหนักตัวอย่าง ดังนั้นหากต้องการปรับค่าถ่วงน้ำหนักตัวอย่างตามการจัดจำแนกใหม่ ต้องลงรหัสใหม่ให้หน่วยสถิติทั้งหมดในประชากร (ทั้งที่สำรวจและไม่ได้สำรวจ) จึงสามารถรวมยอดหน่วยสถิติตัวอย่างตามการจัดจำแนกใหม่ได้

จะเห็นได้ว่าการให้รหัสกิจกรรมของแต่ละหน่วยสถิติที่มาจากข้อมูลที่ละเอียดและน่าเชื่อถือ การใช้วิธีการที่อ้างอิงข้อมูลระดับจุลภาคจะให้ผลที่น่าเชื่อถือมากกว่าผลที่ได้จากการใช้วิธีที่อ้างอิงข้อมูลมหภาคที่จะอธิบายต่อไป

1.2 ประเด็นการจัดจำแนกที่เกี่ยวกับวิธีข้อมูลระดับจุลภาค

วิธีลงรหัสคู่และวิธีข้อมูลระดับจุลภาค

วิธีข้อมูลระดับจุลภาคจำเป็นต้องมีการลงรหัสคู่ตามกิจกรรมหลักของหน่วยสถิติ ซึ่งการลงรหัสคู่ให้กับแต่ละหน่วยสถิติจะทำให้เห็นการเปลี่ยนแปลงของกิจกรรมหลักทั้งในรูปแบบของการจัดจำแนกเดิมและการจัดจำแนกใหม่ ซึ่งสามารถนำการเปลี่ยนแปลงนี้ไปใช้กับทุกช่วงเวลาของอนุกรมที่มีหน่วยสถิตินั้น ๆ อยู่

การลงรหัสคู่สามารถทำได้หลายวิธี วิธีหนึ่งคือสอบถามจากหน่วยสถิติ โดยให้หน่วยสถิติอธิบายลักษณะของกิจกรรมหลักของสถานประกอบการออกมาทั้งตามการจัดจำแนกเดิมและการจัดจำแนกใหม่ ซึ่งในกรณีนี้ควรมีคำอธิบายรายละเอียดให้กับหน่วยสถิติ อย่างไรก็ตามวิธีนี้ไม่จำเป็นต้องกำหนดกิจกรรมหลักให้กับทุกกิจกรรมของสถานประกอบการ โดยเฉพาะในสถานประกอบการที่ดำเนินการหลายกิจกรรม ในการกำหนดกิจกรรมหลักควรใช้วิธีการบน-ล่าง (top-down method)[2] ซึ่งจำเป็นต้องทราบส่วนแบ่งของมูลค่าเพิ่ม (value-added) ของแต่ละกิจกรรมที่ดำเนินการในหน่วยสถิติ หรืออย่างน้อยต้องมีตัวแปรที่เป็นตัวแทนของมูลค่าเพิ่ม ซึ่งถ้าสำนักงานสถิติเป็นผู้ดำเนินการโดยใช้วิธีที่เหมาะสม จะทำให้มีความน่าเชื่อถือมากขึ้น

การใช้การจัดจำแนกตัวกลาง (intermediary classification) ในการลงรหัสกิจกรรมใหม่

เมื่อมีการเปลี่ยนการจัดจำแนกของกิจกรรม การสอบถามแต่ละสถานประกอบการเพื่อให้ได้ตัวเลขมูลค่าเพิ่มที่เกิดจากแต่ละกิจกรรมพื้นฐาน (elementary activity) ของทั้งการจัดจำแนกเดิมและการจัดจำแนกใหม่นั้นเป็นเรื่องยาก ดังนั้นการใช้การจัดจำแนกตัวกลางเพื่อสอบถามเฉพาะข้อมูลที่จำเป็นอาจเป็นทางเลือกที่ดีกว่า

การจัดจำแนกตัวกลางที่ใช้สอบถามมูลค่าเพิ่มของหน่วยธุรกิจ พัฒนาขึ้นเพื่อช่วยกำหนดกิจกรรมหลักของสถานประกอบการทั้งตามการจัดจำแนกเดิมและการจัดจำแนกใหม่ โดยไม่เกิด การซ้ำซ้อนกันของการสอบถาม การจัดจำแนกตัวกลางเป็นสับเซตของผลคูณคาร์ทีเซียนของทั้งสองการจัดจำแนก โดยมีรายละเอียดและตัวอย่าง ดังนี้

สมมติให้การจัดจำแนกเดิมคือ

- หมู่ใหญ่ A และ หมู่ใหญ่ B

[2] ดูที่ International Standard Industrial Classification of All Economic Activities (ISIC) Revision 4, หน้า 23-28

- หมู่ใหญ่ทั้งสองส่วนถูกแบ่งเป็น หมู่ย่อย A1 A2 และ A3 อีกส่วนแบ่งเป็น B1 และ B2

สมมติว่าการจัดจำแนกใหม่ปรับโครงสร้างของหมู่ย่อย ดังนี้

- หมู่ใหญ่ X ประกอบด้วย หมู่ย่อย A1
- หมู่ใหญ่ Y ประกอบด้วย หมู่ย่อย A2 A3 และ B1
- หมู่ใหญ่ Z ประกอบด้วย หมู่ย่อย B2

การจัดจำแนกตัวกลางระหว่างการจัดจำแนกเดิมและการจัดจำแนกใหม่สามารถทำได้ดังนี้[3]

$$M = \{A1\},\ N = \{A2, A3\},\ O = \{B1\},\ P = \{B2\}$$

สามารถกำหนดการจัดจำแนกทั้งเดิมและใหม่ได้โดยตรง โดยใช้การจัดจำแนกตัวกลางดังนี้

$$A = \{M, N\};\ B = \{O, P\}$$

$$X = \{M\};\ Y = \{N, O\};\ Z = \{P\}$$

การสังเกตกิจกรรมพื้นฐานของหน่วยสถิติทั้งในรูปแบบของการจัดจำแนกเดิมและการจัดจำแนกใหม่ (โดยใช้การจัดจำแนกตัวกลาง) เป็นประโยชน์อย่างยิ่งสำหรับการคำนวณย้อนกลับ โดยไม่เพียงแต่สนใจสถิติในระดับ ISIC เท่านั้น (เช่น สถิติที่อ้างอิงตามกิจกรรมหลักของหน่วยสถิติ) แต่ยังสนใจสถิติทุกภาคส่วนในระดับกิจกรรมพื้นฐาน ซึ่งการสังเกตทุกกิจกรรมพื้นฐานทั้งสองการจัดจำแนกจะช่วยให้คำนวณย้อนกลับได้โดยที่กิจกรรมที่เกี่ยวข้องกันถูกคำนวณไปพร้อม ๆ กัน และมีความสอดคล้องกัน (อ้างอิงตามข้อมูลเดียวกัน) ซึ่งสามารถทำได้เนื่องจาก ในข้อมูลระดับสถานประกอบการจะกำหนดกิจกรรมหลักจากการสังเกตกิจกรรมพื้นฐานทุก ๆ กิจกรรมตามการจัดจำแนกใหม่

ข้อดีอีกประการหนึ่งคือ การสังเกตทุกกิจกรรมพื้นฐานจะช่วยให้สมมติฐาน “โครงสร้างคงที่” ของการพยากรณ์ย้อนหลังอ่อนลง ในวิธีการสัดส่วน (ดูที่หัวข้อ 2) จะนำโครงสร้างของการเปลี่ยนแปลงระหว่างการจัดจำแนกเดิมและการจัดจำแนกใหม่ไปใช้กับทุก ๆ ปี ในอนุกรม โดยอ้างอิงตามการสังเกตใน “ปีที่ลงรหัสคู่” (โครงสร้างจึงคงที่) ส่วนในวิธีข้อมูลระดับจุลภาคจะเชื่อถือสมมติฐาน“โครงสร้างคงที่” น้อยกว่า เนื่องจากวิธีนี้ถูกสร้างขึ้นในระดับสถาน

[3] สามารถสร้างการจัดจำแนกตัวกลางได้หลากหลายระดับ เช่น ในระดับที่มีรายละเอียดมากกว่านี้ ซึ่งการจัดจำแนกตัวกลางทั้งหมดที่แสดงเหล่านี้มาจากการจัดจำแนกตัวกลางที่ต้องสังเกต “น้อยที่สุด” (ให้รายละเอียดที่จำเป็นมากที่สุด)

ประกอบการ ไม่ใช่ในระดับข้อมูลรวมยอด ซึ่งถ้า “สมมติฐานโครงสร้างคงที่” (ที่เป็นสมมติฐานในวิธีการคำนวณใด ๆ) มีการสร้างในระดับที่ละเอียดกว่าการจัดจำแนกปกติ จะทำให้สมมติฐานนี้มีความหนักแน่นน้อยลงเช่นกัน อีกทั้งยังสะท้อนให้เห็นถึงการพัฒนาของโครงสร้างเศรษฐกิจในอนุกรมที่คำนวณย้อนกลับได้มากขึ้น

ขั้นตอนที่อธิบายในข้างต้นจะคำนึงถึงการเปลี่ยนแปลงที่เป็นไปได้ทั้งหมด จากองค์ประกอบในแต่ละกิจกรรม ของแต่ละหน่วยสถิติ ในแต่ละปี ตัวอย่างดังต่อไปนี้จะแสดงหลักการและข้อดีของวิธีการนี้

สมมติให้สถานประกอบการ E ตามการจัดจำแนกเดิมประกอบด้วย หมวด A และ B แต่ละหมวดประกอบด้วยหมู่ A1, A2 และ B1, B2 และสมมติให้การจัดจำแนกใหม่แตกต่างจากเดิมคือแบ่ง A2 เป็น 2 หมู่ คือ A21 และ A22 และกำหนดหมวดใหม่ คือ หมวด U ประกอบด้วย A1 และ A21 และหมวด V ประกอบด้วย A22, B1 และ B2 ตารางเปรียบเทียบระหว่างการจัดจำแนกเดิมกับการจัดจำแนกใหม่ แสดงดังต่อไปนี้

การจัดจำแนกเดิม			การจัดจำแนกใหม่
A1		A1	U
A2		A21	
		A22	V
B1		B1	
B2		B2	

พิจารณาส่วนแบ่งมูลค่าเพิ่มของสถานประกอบการ E ในปีที่มีการลงรหัสคู่[4]

ปีที่ T	
A1	10
A21	10
A22	20
B1	20
B2	10

ในปีที่ T ใช้วิธีแบบบนลงล่าง ซึ่งถ้าคำนวณตามการจัดจำแนกเดิม กิจกรรมหลักคือ A (40 ต่อ 30) แต่ถ้าคำนวณตามการการจัดจำแนกใหม่กิจกรรมหลัก คือ V (50 ต่อ 20) ซึ่งหาก

[4] ในตัวอย่างนี้ การลงรหัสคู่จะเทียบเท่ากับการลงรหัสใหม่ เนื่องจากการจัดจำแนกใหม่ประกอบด้วยการแบ่งหมวด A2 เป็น หมู่ A21 และ A22

อ้างอิงตามค่าสังเกตที่ได้จากการลงรหัสคู่ของสถานประกอบการ E ในปีที่ T กิจกรรมพื้นฐาน A2 ถูกแบ่งออกเป็น A21 และ A22 ในสัดส่วน 1/3 และ 2/3 ตามลำดับ

สมมติให้ปีที่ R เป็นปีที่ถูกประมาณย้อนกลับ โดยมีมูลค่าเพิ่มตามการจัดจำแนกเดิมคือ

ปีที่ R	
A1	30
A2	12
B1	10
B2	10

กิจกรรมหลักตามการจัดจำแนกเดิมกิจกรรมหลักของสถานประกอบการ E คือ A (42 ต่อ 20)

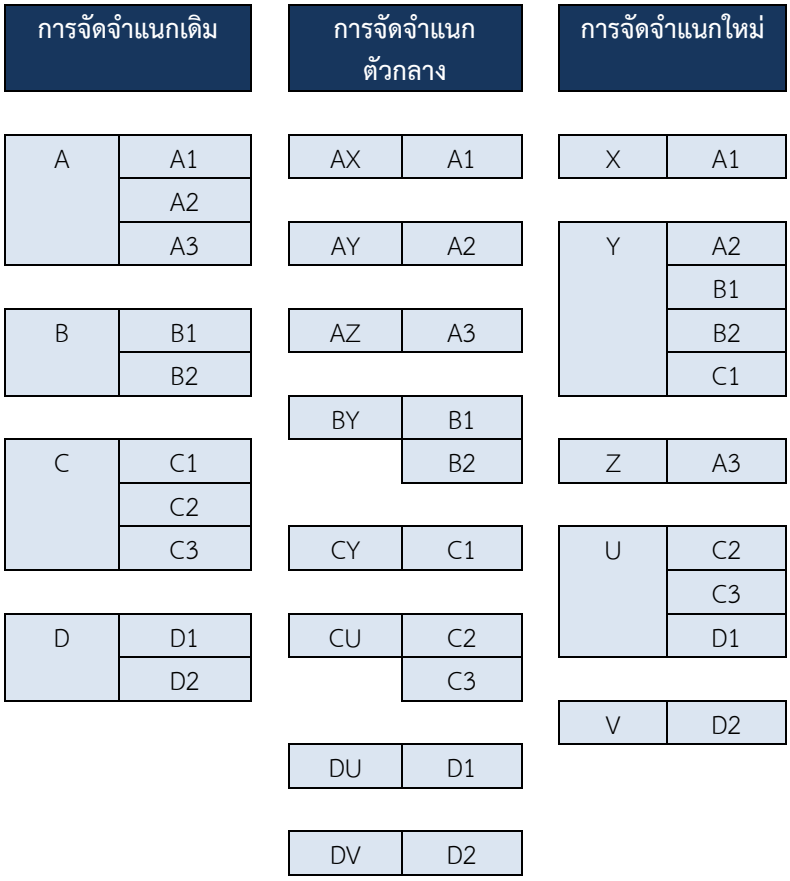
ถ้าใช้วิธีจุลภาคกับกิจกรรมหลักโดยตรง รหัส A (การจัดจำแนกเดิม) ในสถานประกอบการ E ในปีที่ R จะได้รับการแปลงรหัสโดยตรงเป็น รหัส V (การจัดจำแนกใหม่) บนพื้นฐานของ เมทริกซ์การเปลี่ยนแปลงตามค่าสังเกตในปีที่ T แต่ถ้าต้องการดำเนินการในระดับที่ละเอียดกว่ากิจกรรมหลัก (กิจกรรมพื้นฐาน) สามารถกำหนดสัดส่วนของค่าสังเกตในปีที่ T ที่แบ่งกิจกรรม A ออกเป็นกิจกรรม A21 และ A22 ตามสัดส่วน 1/3 และ 2/3 เพื่อประมาณมูลค่าเพิ่มตามการจัดจำแนกใหม่ได้ ดังนี้

ปีที่ R (ประมาณค่า)	
A1	30
A21	4
A22	8
B1	10
B2	10

จากตารางข้างต้น สามารถสรุปได้ว่า “กิจกรรมหลักที่ได้จากการคำนวณย้อนกลับ” คือ U ไม่ใช่ V (34 ต่อ 28) จะเห็นได้ว่าการคำนวณย้อนกลับที่ทำงานในระดับละเอียดมากที่สุด (กิจกรรมพื้นฐานของหน่วยสถิติ) ช่วยให้การปรับเปลี่ยนโครงสร้างของหน่วยสถิติระหว่างปีที่ R และ T มีความน่าเชื่อถือและสะท้อนความเป็นจริงมากยิ่งขึ้น

ตัวอย่างการสร้างการจัดจำแนกตัวกลาง

แผนผังต่อไปนี้แสดงให้เห็นถึง หมวด A, B, C และ D ของการจัดจำแนกแบบเดิมจะเชื่อมโยงกับหมวด X, Y, Z, U และ V ที่จัดจำแนกแบบใหม่



การจัดจำแนกตัวกลางประกอบด้วยรหัส AX, AY, AZ, BY, CY, CU, DU และ DV

การเชื่อมโยงรหัสระหว่างการจัดจำแนกเดิมและการจัดจำแนกใหม่ทำได้โดยผ่านการจัดจำแนกตัวกลาง (ในตัวอย่าง)

การจัดจำแนกตัวกลางเป็นสับเซตของผลคูณคาร์ทีเซียนของการจัดจำแนกเดิมและการจัดจำแนกใหม่ กล่าวคือ เป็นความน่าจะเป็นของการเชื่อมโยงระหว่างสองการจัดจำแนกที่เป็นไปได้ทั้งหมด ดังนั้นจึงไม่มีรหัส AU, AV, BX, BZ, BU, BV, CX, CZ, CV, DX, DY และ DZ เนื่องจากตรงกับเซตว่าง (ไม่มีจุดตัดระหว่างรหัสทั้งสอง)

รหัส A1, A2, ..., D1, D2 เป็นรหัสที่ถูกแยกรายละเอียดจากการจัดจำแนกเดิมหรือการจัดจำแนกใหม่ การแยกรหัสเหล่านี้เกิดจากการแยกรายละเอียดของกิจกรรม หรือผลิตภัณฑ์ หรือกลุ่มของผลิตภัณฑ์ ซึ่งการแยกรายละเอียดตามการจัดจำแนกตัวกลางนั้นดีกว่าการแยกรหัสออกจากกันทุก ๆ รหัส เช่น การแยกรหัส C2 และ C3 ออกจากกันนั้น

ไม่ได้ช่วยให้ได้ข้อมูลเพิ่มเติมมากกว่าเซตของ {C2, C3} รวมถึงรหัส B1 และ B2 เช่นกัน นอกจากนี้จำนวนของรหัสการจัดจำแนกตัวกลาง (8) คือจำนวนที่น้อยที่สุดที่เป็นไปได้ในตัวอย่างนี้ และน้อยกว่าตามทฤษฎี (การเชื่อมโยงที่เป็นไปได้ทั้งหมดคือ 4x5 = 20)

วิธีปฏิบัติต่อหน่วยสถิติที่ไม่ทราบกิจกรรม

อาจเกิดเหตุการณ์ที่ไม่ทราบกิจกรรมของหน่วยสถิติที่ต้องคำนวณย้อนกลับในช่วงเวลาก่อนการลงรหัสคู่ (ปีที่ R) เช่น ไม่ได้เก็บ หน่วยสถิติไม่ถูกสำรวจ หรือไม่ตอบกลับ

มีสถานการณ์ที่เป็นไปได้มีสองสถานการณ์คือ

- หน่วยสถิติถูกสังเกตในปีที่ลงรหัสคู่ แต่ไม่ได้ถูกสังเกตในช่วงเวลาก่อนหน้า ในกรณีนี้วิธีแก้ปัญหาที่ปลอดภัยที่สุดคือ สมมติให้กิจกรรมหลักในปีที่ R เหมือนกับกิจกรรมในปีที่ลงรหัสคู่
- หน่วยสถิติไม่ถูกสังเกตในปีที่ลงรหัสคู่ แต่มีอยู่ในปีที่ R ในกรณีนี้ วิธีการแก้ปัญหาคือ การใช้เมทริกซ์เปลี่ยนสถานะ (transition matrix) เพื่อกำหนดกิจกรรมหลักให้ปีที่ R ตามการจัดจำแนกใหม่ โดยใช้ทั้งวิธีจุลภาค และวิธีมหภาค (ใช้เมทริกซ์การเปลี่ยนแปลง) ร่วมกัน

ใช้วิธีเมทริกซ์เปลี่ยนสถานะ หรือ วิธีค่าใกล้ที่สุด?

วิธีเมทริกซ์เปลี่ยนสถานะ แสดงความน่าจะเป็นที่รหัสกิจกรรมที่ i ตามการจัดจำแนกเดิมจะเปลี่ยนเป็นรหัสตามการจัดจำแนกใหม่ ความน่าจะเป็นนี้กำหนดโดยอ้างอิงจากความถี่ที่ถูกสังเกตของประชากรในปีที่มีลงรหัสคู่

มี “การเปลี่ยนแปลง” สองประเภทที่อาจส่งผลกระทบต่อ class i ตามการจัดจำแนกเดิม[5]

- Class i มีความสัมพันธ์แบบ one-to-one กับ class j ที่เป็นการจัดจำแนกใหม่ (มีหรือไม่มีการเปลี่ยนแปลงของรหัส)
- Class i มีความสัมพันธ์แบบ one-to-many กับ class j (class i 1 class สามารถแบ่งเป็น class j ได้หลาย class)

[5] แม้ว่าระหว่าง class i กับ class j นั้นเหมือนกัน แต่ไม่มีเหตุผลรองรับที่จะสมมติให้หน่วยสถิติที่มีกิจกรรมหลักเป็น class i ตามการจัดจำแนกเดิม เปลี่ยนเป็นกิจกรรมหลัก class j ตามการจัดจำแนกใหม่ได้ โดยจะสามารถสมมติได้ก็ต่อเมื่อหน่วยสถิตินั้นดำเนินกิจกรรมเพียงอย่างเดียว นอกจากนี้การใช้วิธีบน-ล่างอาจส่งผลกระทบต่อการกำหนดกิจกรรมหลักตามการจัดจำแนกใหม่ได้

ในกรณีแรกมีสัมประสิทธิ์การเปลี่ยนแปลงเพียงค่าเดียว ci,j=1 (อื่น ๆ จะเท่ากับ 0) ส่วนในกรณีที่สอง ค่าสัมประสิทธิ์จะอยู่ระหว่าง 0 ถึง 1 และผลรวมทั้งหมดคือ 1 โดยในกรณีที่สอง การใช้เมทริกซ์การเปลี่ยนแปลงเพื่อลงรหัสกิจกรรมใหม่นั้นเป็นการแก้ไขโครงสร้างกิจกรรมของสถานประกอบการอย่างที่ไม่ได้อ้างอิงค่าที่แท้จริง (ความน่าจะเป็น)

นอกจากนี้สามารถประยุกต์วิธีการที่ง่ายขึ้นได้ โดยใช้ “วิธีค่าที่ใกล้ที่สุด” (nearest-neighbor procedure) แทนเมทริกซ์การเปลี่ยนแปลง ซึ่งวิธีนี้ดำเนินการโดยหาหน่วยสถิติที่ “ใกล้เคียงที่สุด” กับหน่วยสถิติที่ไม่ทราบค่า และคำนวณย้อนกลับให้หน่วยสถิตินั้นในลักษณะเดียวกับที่ดำเนินการกับ "หน่วยที่ใกล้เคียงที่สุด"

2 วิธีที่อ้างอิงค่าสัมประสิทธิ์การเปลี่ยนแปลง (วิธีสัดส่วน)

2.1 วิธีสัดส่วนคืออะไร

"วิธีสัดส่วน" คือวิธีการคำนวณย้อนกลับอย่างง่าย ที่พยายามกำหนดโครงสร้างให้กับอนุกรมเวลาใหม่ โดยนำช่วงเวลาชั่วคราว (transitory period) ที่เป็นช่วงเวลาที่หน่วยสถิติทุกหน่วยได้รับการลงรหัสคู่ (ตามการจัดจำแนกเดิมและการจัดจำแนกใหม่) อย่างน้อย 1 ช่วงเวลา มาปรับโครงสร้างของอนุกรมในอดีตให้เป็นไปตามการจัดจำแนกใหม่ กล่าวคือ เป็นการนำชุดของ "ค่าสัมประสิทธิ์การเปลี่ยนแปลง" ที่คำนวณจากสัดส่วนของหมวด/หมู่ตามการจัดจำแนกเดิมที่แปลงเป็นหมวด/หมู่ตามการจัดจำแนกใหม่ มาใช้เป็นสัดส่วนอ้างอิงให้กับอนุกรมในอดีต

การใช้วิธีสัดส่วน ผลลัพธ์ที่ได้ตามการจัดจำแนกเดิมและการจัดจำแนกใหม่นั้นมีความความสัมพันธ์กันสูง เนื่องจากจะกำหนดข้อมูลการจัดจำแนกเดิมโดยอ้างอิงจากสัดส่วนตามกลุ่มของการจัดจำแนกใหม่ ดังนั้นอนุกรมทั้งหมดจะสอดคล้องกัน โดยสมมติให้ความสัมพันธ์ที่ได้จากช่วงเวลาที่ลงรหัสคู่คงที่ตลอดช่วงเวลาที่พยากรณ์ย้อนหลัง

วิธีสัดส่วนสามารถนำไปใช้ได้กับข้อมูลระดับ "มหภาค"[6] เท่านั้น เนื่องจากค่าสัมประสิทธิ์การเปลี่ยนแปลงประมาณค่าจากจำนวนหน่วยตัวอย่าง ดังนั้นจึงไม่จำเป็นต้องเข้าถึงข้อมูลในระดับจุลภาคของหน่วยสถิติ วิธีนี้จึงเป็นวิธีที่ใช้ทรัพยากรในการคำนวณย้อนกลับน้อย ทั้งต้นทุนและเวลา แต่ข้อเสียคือ วิธีนี้เป็นเพียงแค่การประมาณค่าสังเกตก่อนหน้าโดยไม่ได้วิเคราะห์ผลกระทบจากการปรับปรุงอนุกรมเวลาในเชิงลึก

วิธีการอย่างง่ายหรือวิธีที่ซับซ้อน

วิธีสัดส่วนคือ การนำอัตราการเติบโตของอนุกรมเวลาก่อนหน้ามาปรับปรุงให้อยู่ภายใต้การจัดจำแนกใหม่ ซึ่งรูปแบบที่ง่ายที่สุดคือ สมมติให้อัตราการเติบโตของอนุกรมที่คำนวณย้อนกลับนั้นไม่เปลี่ยนแปลงตลอดทั้งอนุกรม อย่างไรก็ตาม มีวิธีที่ซับซ้อนมากขึ้นโดยการกำหนดสัมประสิทธิ์ให้กับแต่ละปีที่ประมาณค่าโดยเฉพาะ ซึ่งสามารถทำได้ตามดุลยพินิจของผู้เชี่ยวชาญหรือบนพื้นฐานของเทคนิคการประมาณค่า

[6] สถิติที่มีการคำนวณ รวมยอด หรือจัดกลุ่มแล้ว เช่น ตารางไขว้ ที่แสดงความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป เป็นต้น

ตัวอย่างที่น่าสนใจคือ กรณีของโทรศัพท์มือถือ จะเห็นได้ว่ารูปแบบการเติบโตในช่วงก่อนการเกิดโทรศัพท์มือถือบางประเภทกับช่วงหลังจากเริ่มขายนั้นมีรูปแบบการเจริญเติบโตที่เฉพาะ ดังนั้นจึงต้องปรับค่าสัมประสิทธิ์การเปลี่ยนแปลง เพื่อให้สอดคล้องกับข้อเท็จจริงที่ทราบนี้

สมมติฐานพื้นฐานของวิธีการสัดส่วน

วิธีสัดส่วนจะทำการแก้ไขเฉพาะค่าประมาณและจะไม่พิจารณาหรือแก้ไขข้อมูลในระดับจุลภาคที่ใช้ในการสร้างค่าประมาณเหล่านี้ ดังนั้นจึงไม่มีการเชื่อมโยงระหว่างกันของข้อมูลระดับจุลภาคและข้อมูลระดับมหภาคในอดีต

การใช้ชุดของสัมประสิทธิ์ชุดเดียวกันในทุกช่วงเวลานั้น อ้างอิงจากสมมติฐานที่ว่า การกระจายตัวของตัวแปรที่สนใจระหว่างการจัดจำแนกเดิมและการจัดจำแนกใหม่นั้นไม่มีการเปลี่ยนแปลง อย่างไรก็ตามสมมติฐานนี้อาจมีปัญหาในกรณีสัดส่วนผลประกอบการของอุตสาหกรรมตาม ISIC Rev.3.1 ที่เปลี่ยนการจัดจำแนกเป็น ISIC Rev.4 บางอุตสาหกรรมนั้นมีความเฉพาะและมีการเปลี่ยนแปลงตลอดเวลา ทำให้วิธีการนี้อาจไม่เหมาะสม

สรุปได้ว่าวิธีสัดส่วนสามารถจัดการกับทุกความสัมพันธ์ระหว่างการจัดจำแนกเดิมและการจัดจำแนกใหม่ เป็นวิธีที่ค่อนข้างง่าย และมีต้นทุนต่ำในการนำไปใช้ แต่ต้องมีสมมติฐานที่กำหนดว่าอนุกรมที่ถูกประมาณย้อนกลับสะท้อนความเป็นจริงได้แม่นยำอย่างไร

เมทริกซ์การเปลี่ยนแปลง

เมทริกซ์การเปลี่ยนแปลง คือ เมทริกซ์ขนาด I x J โดยที่ I คือจำนวนหน่วยในหมู่ตามการจัดจำแนกเดิม และ J คือจำนวนหน่วยในหมู่ตามการจัดจำแนกใหม่ ส่วนประกอบของเมทริกซ์ c_{ij} คือความน่าจะเป็นที่กิจกรรมที่ลงรหัส i ตามการจัดจำแนกเดิม ถูกลงรหัสเป็น j ตามการจัดจำแนกใหม่ ดังนั้น $\sum_i c_{ij} = 1$ และ $\sum_j c_{ij} = 1$

ความน่าจะเป็นเหล่านี้ถูกกำหนดจากการเปลี่ยนแปลงที่ได้จากหน่วยสถิติในปีที่มีการลงรหัสคู่ ข้อมูลที่จำเป็นสำหรับการกำหนดเมทริกซ์การเปลี่ยนแปลงคือ กิจกรรมพื้นฐานและกิจกรรมหลักของหน่วยสถิติ

เพื่อให้มีความแม่นยำขึ้น เราสามารถสร้างเมทริกซ์การเปลี่ยนแปลงที่มีเงื่อนไขได้ โดยพิจารณาจากขนาดของหน่วยสถิติ หรือกิจกรรมหลัก หรือจำนวนของกิจกรรมพื้นฐานที่ดำเนินการโดยหน่วยสถิติ

2.2 การใช้วิธีสัดส่วน

ขั้นตอนในการใช้วิธีสัดส่วนแสดงได้ดังต่อไปนี้

จุดเริ่มต้น – ตารางเปรียบเทียบ

ขั้นตอนแรกของวิธีสัดส่วนคือ การสร้างตารางเปรียบเทียบ เนื่องจากตารางเปรียบเทียบเป็นการสร้างการเชื่อมโยงระหว่างระบบการจัดจำแนกเดิมกับการจัดจำแนกใหม่ และแสดงให้เห็นถึงความสัมพันธ์ของการจัดจำแนกเดิมไปการจัดจำแนกใหม่และการจัดจำแนกใหม่ไปยังการจัดจำแนกเดิม รวมถึงให้ข้อมูลการเปลี่ยนแปลงของทั้งสองระบบ (ส่วนใหญ่เป็นข้อมูลเชิงปริมาณ)

ตารางเหล่านี้จะช่วยให้ผู้ใช้เข้าใจความสัมพันธ์ระหว่างการจัดจำแนกเดิมและการจัดจำแนกใหม่ เห็นถึงขอบเขตของการเปลี่ยนแปลง และเข้าใจว่าการปรับปรุงที่เกิดขึ้นส่งผลกระทบต่อความต่อเนื่องของข้อมูลอย่างไร ในส่วนของผู้ผลิตสถิติ ตารางเปรียบเทียบจะเป็นพื้นฐานสำหรับการคำนวณค่าสัมประสิทธิ์การเปลี่ยนแปลง

ตารางเปรียบเทียบอาจให้รายละเอียดได้มากหรือน้อย ขึ้นอยู่กับคำอธิบายที่ให้รายละเอียดของการเปรียบเทียบระหว่างระบบเดิมและระบบใหม่ สำหรับตารางเปรียบเทียบเพื่อการแปลงข้อมูลจากระบบหนึ่งไปยังอีกระบบหนึ่งอย่างน้อยควรมีรายละเอียดคือ (1) รายการอุตสาหกรรมทั้งหมดในแต่ละหมวดหมู่ และ (2) การเปลี่ยนแปลงขอบเขตของอุตสาหกรรม (การเพิ่มและการลด) จากระบบเดิมไปยังระบบใหม่ และจากระบบใหม่ไประบบเดิม

ขั้นที่ 1 – การประมาณค่าของสัมประสิทธิ์การเปลี่ยนแปลง

สัมประสิทธิ์การเปลี่ยนแปลง คือปัจจัยที่อ้างอิงจากการจัดกลุ่มข้อมูลอุตสาหกรรมใหม่ในระดับรวมยอด ที่สะท้อนให้เห็นการเปลี่ยนแปลงระหว่างระบบการจัดจำแนกเดิมและการจัดจำแนกใหม่ ซึ่งควรคำนวณในระดับที่ละเอียดที่สุดที่เป็นไปได้

การสร้างตารางเปรียบเทียบจะช่วยคำนวณสัมประสิทธิ์การเปลี่ยนแปลงให้แต่ละหมวดของการจัดจำแนกได้ โดยอ้างอิงตามจำนวนหน่วยสถิติ และยังสามารถคำนวณโดยอ้างอิงตามตัวแปรได้ เช่น ผลประกอบการ การจ้างงาน รายได้ การขาย ฯลฯ ซึ่งจำเป็นต้องใช้ข้อมูลระดับจุลภาคเข้ามาช่วยในการคำนวณ นอกจากนี้ยังสามารถใช้ชุดสัมประสิทธิ์แต่ละชุดที่แตกต่างกันตามตัวแปรที่สนใจได้

สัมประสิทธิ์การเปลี่ยนแปลงจะแสดงจำนวนของอุตสาหกรรมแต่ละอุตสาหกรรมว่ามีการเปลี่ยนแปลงเท่าไร (ในรูปของหน่วยสถิติหรือในรูปของตัวแปร) มีการเคลื่อนไหวเกิดที่ใด

เกิดการเปลี่ยนแปลงระหว่างอุตสาหกรรมใด และไปในทิศทางใด หรือเรียกได้ว่าสัมประสิทธิ์การเปลี่ยนแปลงคือการนำเสนอตารางเปรียบเทียบในเชิงปริมาณ

สัมประสิทธิ์การเปลี่ยนแปลงสามารถคำนวณได้โดยอ้างอิงเพียงช่วงเวลาเดียวหรือหลายช่วงเวลา ซึ่งข้อได้เปรียบของการประมาณค่าสัมประสิทธิ์จากหลายช่วงเวลาคือ ช่วยในการตัดสินใจว่าสัมประสิทธิ์การเปลี่ยนแปลงที่คำนวณ ณ ช่วงเวลาเดียวนั้นเหมาะสมหรือไม่ ในทางทฤษฎี การมีค่าสัมประสิทธิ์การเปลี่ยนแปลงสำหรับทุก ๆ ช่วงเวลาในอดีตนั้นเป็นการดำเนินงานที่ดีที่สุดในอุดมคติ ซึ่งในทางปฏิบัติอาจเป็นการใช้ทรัพยากรที่มากเกินไป แต่มีวิธีการที่สามารถทำได้ในทางปฏิบัติคือ การคำนวณสัมประสิทธิ์โดยใช้ 2 ช่วงเวลาที่ต่างกัน (เช่น ช่วงเวลา ณ จุดเริ่มต้นและจุดสิ้นสุดของอนุกรมในอดีต) จากนั้นใช้วิธีการประมาณค่าในช่วง (interpolation) ในการคำนวณค่าสัมประสิทธิ์ระหว่างสองช่วงเวลาดังกล่าว

การคำนวณค่าสัมประสิทธิ์การเปลี่ยนแปลงสามารถคำนวณได้โดยอ้างอิงข้อมูลอย่างน้อย 1 ปี (หรือไตรมาส หรือเดือน) ซึ่งเป็นปี (หรือไตรมาส หรือเดือน) ที่เปลี่ยนแปลงจากการจัดจำแนกหนึ่งไปยังการจัดจำแนกหนึ่ง สำหรับการปรับปรุงคุณภาพของสัมประสิทธิ์การเปลี่ยนแปลง แนะนำให้ขยายระยะเวลาของการลงรหัสคู่ ตัวอย่างเช่น เพิ่มการลงรหัสคู่อีก 1 ปี (หรือไตรมาส หรือเดือน) เพื่อให้ (1) ระยะเวลาในการจัดจำแนกใหม่คงที่ (2) มีสัมประสิทธิ์ที่คำนวณโดยอ้างอิงจากข้อมูลที่ได้รับการตรวจแก้แล้วบางส่วน และ (3) เพื่อให้มีชุดของสัมประสิทธิ์ที่มีเสถียรภาพมากขึ้นจากการรวมกันของข้อมูลที่ลงรหัสคู่

ขั้นที่ 2 ใช้สัมประสิทธิ์การเปลี่ยนแปลงกับค่าประมาณจากการจัดจำแนกเดิม

ในขั้นตอนที่ 2 สามารถคำนวณค่าประมาณอุตสาหกรรมตามการจัดจำแนกใหม่ได้โดยถ่วงน้ำหนักให้ค่าประมาณอุตสาหกรรมตามการจัดจำแนกเดิม โดยใช้สัมประสิทธิ์การเปลี่ยนแปลงเป็นค่าถ่วงน้ำหนัก จากนั้นทำการหาผลรวม ซึ่งผลลัพธ์ที่ได้จะเท่ากับค่าประมาณตามการจัดจำแนกใหม่ เช่น ให้อุตสาหกรรม A ตามการจัดจำแนกใหม่ประกอบด้วย 2 ส่วนที่มาจาก 2 อุตสาหกรรมที่ต่างกัน คือ A1 และ A2 ตามการจัดจำแนกเดิม ค่าสัมประสิทธิ์การเปลี่ยนแปลง คือค่าสัดส่วนของ A1 และ A2 ในอุตสาหกรรมใหม่ A ตามลำดับ ซึ่งจะแสดงรายละเอียดในตัวอย่างต่อไป

ในบางครั้ง (เมื่อค่าสัมประสิทธิ์การเปลี่ยนแปลง เพียงชุดเดียวถูกนำไปใช้กับอนุกรมเวลาทั้งหมด) การคำนวณนี้เรียกว่า “สมการเชิงเส้นถ่วงน้ำหนัก” (Weight linear combination)

ขั้นตอนที่ 3 การเชื่อมโยงของค่าประมาณทั้งสามช่วงเวลา

วัตถุประสงค์โดยรวมของการพยากรณ์ย้อนหลังคือ เพื่อสร้างชุดอนุกรมในอดีตให้จัดจำแนกตามการจัดจำแนกใหม่ โดยใช้อนุกรมตามการจัดจำแนกเดิมที่มีอยู่ สามารถแบ่งอนุกรม “ตามลำดับเหตุการณ์” ได้ 3 ส่วน คือ

1. ช่วงเวลาในอดีต (historical) ที่มีเพียงการจัดจำแนกเดิมเท่านั้น ส่วนนี้จะเป็นส่วนที่มีการประมาณด้วยค่าสัมประสิทธิ์การเปลี่ยนแปลง
2. ช่วงเวลาชั่วคราว (transitory) คือช่วงเวลาที่นำเสนอทั้งการจัดจำแนกเดิมและการจัดจำแนกใหม่ ซึ่งในช่วงเวลานี้เป็นช่วงที่มีการ “สังเกต” ค่าสัมประสิทธิ์การเปลี่ยนแปลง
3. ช่วงเวลาสุดท้าย (final) เป็นช่วงที่จะใช้เพียงการจัดจำแนกใหม่เท่านั้น

ไม่ว่าจะใช้วิธีการใดเพื่อให้ได้ค่าประมาณสำหรับช่วงเวลาในอดีต มักเกิดการแบ่งช่วงหรือรอยต่อระหว่างส่วนแรก (historical) กับส่วนที่สอง (transitory) ซึ่งการแบ่งช่วงนี้เกิดจากการที่ค่าสังเกตมีการเปลี่ยนแปลง (จากการเก็บรวบรวมตามการจัดจำแนกเดิม เป็นการเก็บรวบรวมโดยลงทั้งรหัสเก่าและใหม่)

สำหรับข้อแนะนำในการลดรอยต่อระหว่างช่วงเวลาในวิธีการแรกคือ การยกระดับให้ช่วงเวลาในอดีตที่ถูกแปลงเทียบเท่ากับช่วงเวลาชั่วคราว เช่น ใช้ปัจจัยหรือตัวแปรของชุดข้อมูลในช่วงเวลาชั่วคราวกับชุดข้อมูลในอดีตตัวแปรเดียวกันในการประมาณค่า ซึ่งวิธีนี้จะช่วยลดรอยต่อระหว่างช่วงเวลาลง อีกวิธีการหนึ่งคือการ “กระจาย” ส่วนที่จะเกิดรอยต่อออก เช่น การแบ่งชุดข้อมูลออกเป็นจำนวนเดือน หรือปี โดยที่ตัวแปรอื่น ๆ ยังคงอยู่

ในขั้นตอนนี้จำเป็นต้องมีนักวิเคราะห์ที่เชี่ยวชาญในโครงการ เพื่อทำการพิจารณาและทบทวนอนุกรมที่ได้ รวมถึงปรับปรุงให้สอดคล้องกับความรู้และประสบการณ์ของพวกเขา

ขั้นตอนที่ 4 ปรับครั้งสุดท้ายเพื่อความสอดคล้อง

เมื่อทำการจัดทำตารางข้อมูลตามการจัดจำแนกใหม่ตามอนุกรมเวลาอย่างต่อเนื่องแล้ว อาจจำเป็นต้องมีการทบทวนและปรับปรุงอนุกรม เพื่อให้แน่ใจว่าข้อมูลรวมยอดตามการจัดจำแนกใหม่สามารถเข้ากับข้อมูลในช่วงอื่น ๆ ได้

ขั้นตอนที่ 5 การปรับฤดูกาล

หนึ่งในวัตถุประสงค์หลักของการพยากรณ์ย้อนหลังคือ เพื่อสร้างอนุกรมเวลาในอดีตและนำมาเป็นพื้นฐานสำหรับการปรับฤดูกาล แต่ขั้นตอนที่อธิบายไว้ข้างต้นอาจไม่ได้กล่าวถึงหรือปรับฤดูกาลของอนุกรมโดยตรง

3 ข้อดีและข้อเสียของวิธีข้อมูลระดับจุลภาคและวิธีสัดส่วน

ข้อดีของ “วิธีข้อมูลระดับจุลภาค”

ข้อดีของวิธีข้อมูลระดับจุลภาคคือ วิธีการนี้สามารถรักษาพัฒนาการทางโครงสร้างของเศรษฐกิจได้ดีที่สุด เนื่องจากวิธีการสัดส่วนดำเนินการในระดับข้อมูลรวมยอด คือนำค่าเมทริกซ์การเปลี่ยนแปลงไปใช้กับแต่ละหมวด และแต่ละปีที่ประมาณย้อนกลับ ตัวอย่างเช่น ถ้าในปีที่ t หมวด S สอดคล้องกับ หมู่ x และหมู่ y ตามการจัดจำแนกใหม่ ซึ่งมีสัดส่วน 30% และ 70% ตามลำดับ จะนำสัดส่วนนี้ไปใช้กับปีที่ถูกประมาณย้อนกลับทั้งหมด เนื่องจากวิธีนี้สมมติให้หน่วยสถิติที่จัดจำแนกไว้ในหมวดต่าง ๆ ในแต่ละปีมีลักษณะเฉพาะที่เหมือนกัน ดังนั้นสัดส่วน 70/30 ที่ได้จากปีที่ t นั้นจะไม่เปลี่ยนแปลงตลอดช่วงเวลา ซึ่งในวิธีข้อมูลระดับจุลภาคไม่จำเป็นต้องใช้สมมติฐานนี้ เนื่องจากแต่ละหน่วยสถิติจะถูกจัดจำแนกใหม่ตามกิจกรรมหลักของสถานประกอบการนั้น ๆ ในแต่ละปีโดยตรง

ข้อดีอีกข้อของวิธีข้อมูลระดับจุลภาคคือ ไม่ต้องเลือกตัวแปรเพื่อใช้ในการคำนวณสัดส่วน เนื่องจากการใช้วิธีสัดส่วนนั้น จำเป็นต้องเลือกตัวแปรอ้างอิงในการกำหนดปัจจัยการเปลี่ยนแปลงหรือสัมประสิทธิ์การเปลี่ยนแปลง เพื่อนำปัจจัยนี้ไปใช้ในการประมาณย้อนกลับ ซึ่งส่วนใหญ่จะใช้จำนวนสถานประกอบการ มูลค่าเพิ่ม หรือจำนวนพนักงาน ดังนั้นสัดส่วนที่นำมากำหนดเมทริกซ์การเปลี่ยนแปลงจึงเป็นเพียงโครงสร้างของตัวแปรที่เลือกเท่านั้น ในขณะที่ตัวแปรอื่น ๆ ที่เกี่ยวข้อง (เช่น ผลประกอบการ เป็นต้น) อาจมีโครงสร้างที่แตกต่างกันออกไปอย่างสิ้นเชิง ตัวอย่างต่อไปนี้แสดงการใช้ตัวแปรที่ต่างกันในการกำหนดเมทริกซ์การเปลี่ยนแปลง และได้ผลที่แตกต่างกัน

สมมติให้หมวด S (ตามการจัดจำแนกเดิม) ประกอบด้วย หน่วยสถิติ U1 U2 และ U3 และมีตัวเลขดังต่อไปนี้

หมวด S	ผลประกอบการ	ลูกจ้าง	มูลค่าเพิ่ม
U1	100	1500	20
U2	200	2400	45
U3	150	2000	50
ผลรวม หมวด S	**450**	**5900**	**115**

ในการจัดจำแนกใหม่ หมวด S (ตามการจัดจำแนกเดิม) ถูกแบ่งออกเป็น หมวด S1 และ S2 โดยที่หมวด S1 ประกอบด้วยหน่วยสถิติ U1 และ U2 ส่วนหมวด S2 ประกอบด้วยหน่วยสถิติ U3

ในกรณีนี้ใช้ “มูลค่าเพิ่ม” เป็นตัวแปรอ้างอิงในวิธีสัดส่วน ดังนั้นจะได้ข้อมูลตามการจัดจำแนกใหม่ ดังนี้

	ผลประกอบการ	ลูกจ้าง	มูลค่าเพิ่ม
หมวด S1	254	3,335	65
หมวด S2	196	2,565	50

แต่วิธีจุลภาคให้ข้อมูล ดังนี้

	ผลประกอบการ	ลูกจ้าง	มูลค่าเพิ่ม
หมวด S1	300	3,900	65
หมวด S2	150	2,000	50

วิธีสัดส่วน[7] พิจารณาจาก หมวด S (การจัดจำแนกเดิม) แบ่งออกเป็น 2 ส่วนคือ S1 และ S2 (การจัดจำแนกใหม่) ตามสัดส่วน 65/115 และ 50/115 โดยสัดส่วนเหล่านี้มาจากมูลค่าเพิ่มที่ได้จากหน่วยสถิติ และนำสัดส่วนเดียวกันนี้ไปใช้กับตัวแปรอื่น ๆ ซึ่งอาจจะไม่ตรงกับความเป็นจริง ส่วนวิธีข้อมูลระดับจุลภาคดำเนินการในระดับหน่วยสถิติจึงไม่จำเป็นต้องใช้สมมติฐานนี้

ข้อเสียของวิธีข้อมูลระดับจุลภาค

ข้อเสียหลักของวิธีข้อมูลระดับจุลภาคคือค่าใช้จ่าย เนื่องจากวิธีข้อมูลระดับจุลภาคต้องทำการลงรหัสให้แต่ละหน่วยสถิติใหม่ในแต่ละช่วงเวลา ทั้งในอนุกรมและระหว่างอนุกรม จากนั้นต้องคำนวณใหม่ในระดับรวมยอด ซึ่งกระบวนการเหล่านี้ส่งผลให้ค่าใช้จ่ายในวิธีข้อมูลระดับจุลภาคสูงกว่าการดำเนินงานกับข้อมูลรวมยอด อย่างไรก็ตามถ้าระยะเวลาของอนุกรมนั้นไม่ยาวมาก อาจมีความเหมาะสมสำหรับค่าใช้จ่ายในการดำเนินงานของวิธีการนี้

นอกจากนี้การใช้วิธีข้อมูลระดับจุลภาคจำเป็นต้องมีข้อมูลกิจกรรมทางเศรษฐกิจของหน่วยสถิติที่ถูกสังเกตในอนุกรมที่สมบูรณ์ เพื่อให้สำนักงานสถิติแห่งชาติสามารถลงรหัสกิจกรรมหลักให้แก่หน่วยสถิติใหม่ โดยอ้างอิงจากรายละเอียดของค่าสังเกตกิจกรรมของหน่วยสถิติ

[7] ตัวอย่างนี้อ้างอิงตามวิธีการมหภาคอย่างง่าย (อ้างอิงตามตัวแปรเพียงตัวเดียว) มีวิธีการมหภาคอื่น ๆ ที่ใช้ประโยชน์จากการรวมตัวแปรเข้าด้วยกัน อย่างไรก็ตามข้อเสียของวิธีสัดส่วนสามารถยังคงเห็นได้ในกรณีนี้

ข้อดีของวิธีสัดส่วน

ข้อดีของวิธีสัดส่วนคือ วิธีสัดส่วนใช้กับข้อมูลระดับ “มหภาค” จึงไม่จำเป็นต้องเข้าถึงข้อมูลในระดับจุลภาคของแต่ละหน่วยสถิติ ดังนั้นจึงใช้ทรัพยากรและเวลาในการคำนวณย้อนกลับน้อย แต่วิธีการนี้เป็นเพียงการแก้ปัญหาโดยประมาณค่าคร่าว ๆ ซึ่งไม่ได้วิเคราะห์ในเชิงลึกถึงผลกระทบของการปรับปรุงอนุกรมเวลา

ข้อเสียของวิธีสัดส่วน

การนำค่าสัมประสิทธิ์ไปใช้กับข้อมูลที่จัดจำแนกตามระบบเดิมเพื่อที่จะแปลงข้อมูลเหล่านั้นให้เป็นไปตามมาตรฐานใหม่ เป็นเพียงแค่การประมาณค่าให้กับค่าสังเกตในอดีต

การที่มีค่าสัมประสิทธิ์การเปลี่ยนแปลงเพื่อนำไปใช้กับทุกช่วงเวลาทั้งช่วงเวลาในอดีตและช่วงเวลาชั่วคราว และสำหรับตัวแปรทุกตัวที่เกี่ยวข้องนั้นเป็นเป็นสิ่งที่ดีและอยู่ในอุดมคติ แต่ในความเป็นจริงอาจมีทรัพยากรที่จำกัด จึงมักจะใช้สัมประสิทธิ์การเปลี่ยนแปลงเพียงหนึ่งชุด ที่สังเกตเป็นเวลาหนึ่งปี (หรือช่วงเวลาอ้างอิงใด ๆ) และคำนวณโดยใช้ตัวแปรอ้างอิงเพียงตัวเดียว (เช่น การจ้างงาน) และนำไปปรับใช้กับตัวแปรอื่น (เช่น รายได้) คำนวณโดยการใช้สัมประสิทธิ์เหล่านี้เป็นตัวถ่วงน้ำหนักคงที่ (fixed weight) ให้กับค่าประมาณข้อมูลในอดีต ซึ่งเป็นการประมาณค่าโดยอ้างอิงตามช่วงเวลาที่เลือก

การดำเนินงานด้วยวิธีนี้อาจเหมาะสมสำหรับช่วงเวลาสั้น ๆ เนื่องจากสมมติฐานที่ประมาณค่าโดยอ้างอิงตามค่าสัมประสิทธิ์อาจก่อให้เกิดความผิดพลาดเป็นระยะเวลานานได้ ถ้าโครงสร้างอนุกรมตามระบบใหม่นั้นแตกต่างจากระบบเดิม

อย่างไรก็ตาม สามารถสร้างสัมประสิทธิ์การเปลี่ยนแปลงได้หลายชุด ตามจำนวนปี “มาตรฐาน” เพื่อที่จะสามารถกำหนดได้ว่าสัมประสิทธิ์การเปลี่ยนแปลงในช่วงเวลาที่เลือกนั้นมีความเหมาะสมหรือไม่

นอกจากนี้การใช้ปีมาตรฐาน หรือดุลยพินิจของผู้เชี่ยวชาญนั้นสามารถช่วยปรับค่าสัมประสิทธิ์การเปลี่ยนแปลงในบางปีที่มีเหตุการณ์สำคัญหรือมีความเฉพาะเจาะจง (เช่น ปรับให้ก่อน ค.ศ. 1990 มีสัมประสิทธิ์ของพนักงานขาย DVD ลดลง) ส่วนค่าสัมประสิทธิ์ของอุตสาหกรรมที่สำคัญที่มีลักษณะเฉพาะ อาจปรับตามเทคนิคในระดับจุลภาค

นอกจากนี้ ข้อเสียของวิธีมหภาคคือ วิธีนี้ไม่ได้แก้ปัญหาในส่วนของกลุ่มอุตสาหกรรมที่มีการเปลี่ยนแปลงเคลื่อนย้ายในตลอดเวลา (ปิดกิจการหรือเปิดกิจการใหม่) โดยเฉพาะปัญหา“หน่วยสถิติก่อนหน้าที่ไม่อยู่ในขอบเขต”

อีกทั้งถ้ามีการใช้ความสอดคล้อง (สัดส่วน) เดียวกันทุก ๆ เดือน ใน 1 ปี จะทำให้รูปแบบตามฤดูกาลของอนุกรมในอดีตบิดเบือน

การปรับปรุงระบบการจัดจำแนกทางเศรษฐกิจมักจะครอบคลุมการเปลี่ยนแปลงที่สะท้อนถึงนวัตกรรมทางด้านอุตสาหกรรม ซึ่งหลักการจัดจำแนกใหม่ของวิธีการนี้อาจไม่สะท้อนถึงความเป็นจริงทางเศรษฐกิจของข้อมูลในอดีต

4 การใช้วิธีประมาณค่าในช่วงระหว่างปีมาตรฐาน (การรวมวิธีข้อมูลระดับจุลภาคและวิธีสัดส่วน)

จากวิธีการที่กล่าวข้างต้น ในทางปฏิบัติมีเพียงส่วนน้อยที่ใช้เทคนิคข้อมูลระดับจุลภาคหรือวิธีสัดส่วนเพียงอย่างใดอย่างหนึ่ง เพราะการเปลี่ยนโครงสร้างอนุกรมเวลาไม่ได้ทำได้เฉพาะในระดับจุลภาคหรือมหภาคเท่านั้น แต่ยังสามารถทำได้โดยการรวมวิธีการทั้งสองเข้าด้วยกัน ซึ่งในหัวข้อนี้จะนำเสนอวิธีการที่รวมวิธีสัดส่วนและวิธีข้อมูลระดับจุลภาคเข้าด้วยกัน เพื่อปรับปรุงคุณภาพของอนุกรมที่พยากรณ์ย้อนหลังเมื่อใช้วิธีการข้อมูลระดับจุลภาคหรือวิธีมหภาคเพียงอย่างเดียว หรือเพื่อลดผลกระทบจากข้อเสียของระบบใดระบบหนึ่ง

ขั้นตอนแรกของวิธีการรวมสามารถอธิบายได้โดยการใช้ ปีมาตรฐาน/การประมาณในช่วง โดยคำนวณค่าสัมประสิทธิ์การเปลี่ยนแปลงจากสองช่วงเวลาที่ต่างกันด้วยวิธีข้อมูลระดับจุลภาค จากนั้นจะนำค่าสัมประสิทธิ์ที่ได้จากสองช่วงเวลามาประมาณค่าในช่วงเพื่อหาค่าสัมประสิทธิ์การเปลี่ยนแปลงให้กับปีอื่น ๆ ที่อยู่ระหว่างสองช่วงเวลานี้ จากนั้นนำค่าสัมประสิทธิ์ที่ได้หาค่าประมาณตามการจัดจำแนกใหม่ด้วยวิธีสัดส่วน ซึ่งวิธีนี้จะลดสมมติฐานจากการอ้างอิงสัมประสิทธิ์ชุดเดียวลง แต่ข้อจำกัดคือจำเป็นต้องมีข้อมูลระดับจุลภาคในช่วงเวลาเดิมที่ต้องการประมาณค่า

ในวิธีการนี้ สองช่วงเวลาที่กล่าวถึงในข้างต้นหรือปีที่มีการลงรหัสคู่นั้นเรียกว่าช่วงเวลามาตรฐาน หรือช่วงเวลาที่จะใช้เป็นปีอ้างอิง ซึ่งการเลือกปีมาตรฐานที่เหมาะสมจะกำหนดโดยผู้เชี่ยวชาญของโครงการ โดยค่าที่ได้จากสัมประสิทธิ์การเปลี่ยนแปลง ทั้ง 2 ชุด จะนำมาใช้แปลงค่าประมาณจากการจัดจำแนกเดิมเป็นการจัดจำแนกใหม่ในระดับรวมยอด ด้วยวิธีการประมาณค่าในช่วง ให้กับอนุกรมเวลาในระหว่างปีมาตรฐานนนี้ ซึ่งอาจมีบางหมู่ในระหว่างปีมาตรฐานที่ไม่ได้มีวิวัฒนาการเป็นเส้นตรง ดังนั้นสามารถใช้วิธี non-linear interpolating ได้

อีกวิธีการที่สามารถทำได้คือ การรวมสัมประสิทธิ์จากแต่ละปีมาตรฐาน เป็น 1 ชุด (ค่าเฉลี่ยของทั้งสองช่วง) จากนั้นนำค่าเฉลี่ยสัมประสิทธิ์การเปลี่ยนแปลงที่ได้ไปใช้กับช่วงเวลาที่ต้องการประมาณค่า วิธีการนี้ยังคงเป็นการประมาณค่าอย่างหยาบ แต่ช่วยให้คุณภาพของการประมาณค่าดีขึ้นกว่าการใช้วิธีสัดส่วนอย่างง่าย

วิธีการสุดท้ายคือ การใช้วิธีข้อมูลระดับจุลภาคในช่วงเวลาที่ต่อเนื่องกัน เช่น สถิติรายเดือนหรือรายไตรมาสที่มีการลงรหัสคู่ต่อเนื่องกัน 3 ปี ทำให้สามารถคำนวณค่าสัมประสิทธิ์การเปลี่ยนแปลงได้ทั้ง 36 เดือน หรือ 12 ไตรมาส ดังนั้นจึงสามารถวิเคราะห์รูปแบบโครงสร้างการเปลี่ยนแปลงได้ โดยสามารถสร้างเป็นแบบจำลองอย่างง่าย ที่อธิบายถึงพัฒนาของสัมประสิทธิ์การ

เปลี่ยนแปลงในช่วงเวลาดังกล่าว ซึ่งแบบจำลองสัดส่วนที่ได้จะช่วยปรับปรุงชุดอนุกรมเวลาในอดีตให้มีคุณภาพดีขึ้น อีกทั้งยังสามารถนำมาใช้ได้กับกรณีที่ช่วงเวลาก่อนหน้านี้ไม่มีข้อมูลระดับจุลภาค

5 ตัวอย่าง

ตัวอย่างที่ 1 การปรับฐานดัชนีอุตสาหกรรมการผลิต ค.ศ. 1991 สถิติเยอรมนี

ที่มา

ในเดือนมกราคม ค.ศ. 1995 การจัดประเภทกิจกรรมทางเศรษฐกิจ ฉบับ 1979 (the Industrial Classification of Economic Activities 1979 Edition: SYPRO)[8] ถูกเปลี่ยนไปใช้ฉบับใหม่คือ WZ 93 (ระดับ 5 หลัก) ซึ่งสามารถเปรียบเทียบกันได้ในระดับ 4 หลัก (Class) กับการจัดจำแนก NACE Rev. 1.1 เนื่องจากมีการเปลี่ยนแปลงการจัดจำแนกทำให้ต้องคำนวณข้อมูลใหม่ โดยข้อมูลที่นำมาใช้ในการคำนวณคือ ข้อมูลอุตสาหกรรมรายเดือนตามการจัดจำแนกเดิม SYPRO และข้อมูลอุตสาหกรรมรายเดือนตามการจัดจำแนกใหม่ WZ93 ตั้งแต่ ค.ศ. 1991 - 1994 รวมถึงข้อมูลผลิตภัณฑ์ตามการจัดจำแนกประเภทผลิตภัณฑ์สำหรับผลิตสถิติ 1989 (the 1989 Product Classification for Production Statistics: GP 89)[9]

สถิติเยอรมนีเลือกใช้วิธีการมหภาคในการประมาณค่าข้อมูลย้อนหลังเนื่องจากเป็นวิธีการที่ง่ายและใช้ระเวลาสั้น ๆ ในการดำเนินการ อีกทั้งการพยากรณ์ย้อนหลังที่ต้องเข้าถึงข้อมูลระดับจุลภาคนั้นมีข้อจำกัดจึงไม่สามารถทำได้

การคำนวณการจัดสรรปัจจัย

สำหรับการแปลงโครงสร้างข้อมูลรายเดือนตามการจัดจำแนก SYPRO ให้เป็น WZ 93 ทำได้โดยเชื่อมโยงแต่ละผลิตภัณฑ์ตามการจัดจำแนก GP 89 กับกิจกรรมทางเศรษฐกิจตาม WZ 93 ในระดับ heading และเชื่อมโยงการจัดจำแนก SYPRO กับการจัดจำแนกผลิตภัณฑ์ GP 89 เช่นกัน จะเห็นได้ว่าสามารถคำนวณการแปลงข้อมูลตาม SYPRO ไปเป็นข้อมูลตาม WZ 93 ได้จากมูลค่าการผลิตรวมตาม GP 89

[8] Systematik der Wirtschaftszweige (Ausgabe 1979), Fassung für die Statistik im Produzierenden Gewerbe – Industrial Classification of Economic Activities (1979 Edition), Version for Statistics of Production Industries.

[9] Klassifikation der Wirtschaftszweige – Industrial Classification of Economic Activities (1993 Edition).

ตาราง 1 วิธีการคำนวณการจัดสรรปัจจัยสำหรับ SYPRO

S_j	GP_{ij}	BPW_{ij} (€)	BPW_j (€)	A_{ij}	W_K สัดส่วน	W_K	BPW_K
S_1	GP_{11}	150		0.3	W_1		
	GP_{21}	100		0.2	W_3	W_1	650
	GP_{31}	250		0.5	W_2		
รวม S_1	GP	500	500	1.0		W_2	350
S_2	GP_{12}	100		0.1	W_2		
	GP_{22}	400		0.4	W_3	W_3	500
	GP_{32}	500		0.5	W_1		
รวม S_2	GP	1000	1000	1.0			

โดยที่

Sj = กิจกรรมทางเศรษฐกิจตาม SYPRO class j (ระดับ Heading 4 หลัก)

Wk = กิจกรรมทางเศรษฐกิจตาม WZ 93 class k (ระดับ Heading 4 หลัก)

GPij = ผลิตภัณฑ์ i ที่จัดตามกิจกรรมทางเศรษฐกิจ class j (สอดคล้องกับ GP 89)

BPWij = มูลค่าการผลิตรวมของ GP ระดับ Headings ที่จัดตาม class j (มาร์คเยอรมัน)

BPWj = มูลค่าการผลิตรวมตาม SYPRO class j

BPWk = มูลค่าการผลิตรวมของกิจกรรมทางเศรษฐกิจ class k (WZ 93)

Aij = ปัจจัยสำหรับการกำหนดมูลค่าการผลิตรวมของ SYPRO ไปยัง WZ 93

ปัจจัย A_{ij} ถูกนำมาใช้สำหรับคำนวณค่าสัมบูรณ์ทั้งหมดในการคำนวณดัชนี ซึ่งในกรณีการคำนวณดัชนีการผลิต ข้อมูลที่เกี่ยวกับมูลค่าเพิ่มเหล่านี้จำเป็นต้องใช้ในการถ่วงน้ำหนัก

การคำนวณปัจจัย การเปลี่ยนแปลง

สำหรับการแปลงดัชนีกิจกรรมทางเศรษฐกิจตาม SYPRO ในแต่ละ class j ไปยังดัชนีกิจกรรมทางเศรษฐกิจตาม WZ 93 ในแต่ละ class k ต้องมีปัจจัยการเปลี่ยนแปลงที่ได้สัดส่วนกัน คือ U_{jk} โดยวิธีการคำนวณ U_{jk} แสดงดังตาราง 2

ตาราง 2 การสร้างปัจจัยการเปลี่ยนแปลงของ SYPRO

W_k	GP_{ij}	BPW_{ij} (DM)	สัดส่วน S_j	U_{jk}
W_1	GP_{11}	150	S_1	0.231
	GP_{32}	500	S_2	0.769
รวม W_1		*650*		*1.000*
W_2	GP_{31}	250	S_1	0.714
	GP_{12}	100	S_2	0.286
รวม W_2		*350*		*1.000*
W_3	GP_{21}	100	S_1	0.200
	GP_{22}	400	S_2	0.800
รวม W_3		*500*		*1.000*

มูลค่าการผลิตทั้งหมดในแต่ละ class ของกิจกรรมทางเศรษฐกิจตาม WZ 93 ประกอบด้วยมูลค่าการผลิตของ SYPRO จากหลาย class ดังนั้นสัดส่วนในแต่ละ class คือ “โครงสร้างการถ่วงน้ำหนัก” ซึ่งจะนำไปคำนวณกับ SYPRO ระดับข้อมูลรวมยอด ในแต่ละ class เพื่อสร้าง class ตามการจัดจำแนกใหม่

ตัวอย่างตาราง 3 แสดงการคำนวณปัจจัยการเปลี่ยนแปลงของ SYPRO โดยมี class 3324 (การผลิตจักรยาน) และ class 3327 (การผลิตชิ้นส่วนจักรยานยนต์และจักรยาน) ที่กำหนดให้เป็นไปตามการจัดจำแนกใหม่ WZ 93 class 35.42 (การผลิตจักรยาน)

ตาราง 3 ตัวอย่างการคำนวณปัจจัยการเปลี่ยนแปลง

Class ของกิจกรรมทางเศรษฐกิจ		มูลค่าการผลิต class ของกิจกรรมทางเศรษฐกิจ 35.42	ปัจจัยการเปลี่ยนแปลง
WZ 93	**SYPRO**	**1,000 มาร์คเยอรมัน**	**U_{jk}**
35.42 การผลิตจักรยาน	3324 การผลิตจักรยาน	1,059,322	0.669
	3327 การผลิตชิ้นส่วนจักรยานยนต์และจักรยาน	524,115	0.331
ทั้งหมด		***1,583437***	***1.000***

SYPRO class 3324 มีมูลค่าการผลิตทั้งหมด 1.06 พันล้านมาร์คเยอรมัน และ SYPRO class 3327 มีมูลค่าการผลิตทั้งหมด 0.5 พันล้าน ทั้งสอง class ถูกจัดสรรให้อยู่ใน WZ 93 class

35.42 ดังนั้นแนวโน้มของดัชนีการผลิตตาม WZ 93 รหัส 35.42 ระดับ 4 หลัก (heading) แสดงได้โดยการรวมกันของดัชนีในแต่ละกิจกรรมทางเศรษฐกิจตาม SYPRO class 3324 และ 3327 โดยใช้ปัจจัยการเปลี่ยนแปลง U_{jk}

การนำปัจจัยการเปลี่ยนแปลงไปใช้

ก่อนที่จะสร้างอนุกรมในระยะยาว ต้องปรับฐานดัชนี SYPRO ค.ศ. 1991=100

ตาราง 4 การนำปัจจัยการเปลี่ยนแปลงไปใช้ (1) ปรับฐานดัชนี SYPRO

กิจกรรมทางเศรษฐกิจ SYPRO	ช่วงเวลา	ดัชนีการผลิต (เดือน)	
		1985=100	1991=100 (ปรับฐาน)
3324 การผลิตจักรยาน	1988	110.3	70.1
	1989	136.0	86.4
	1990	161.2	102.4
	1991	157.4	100.0
3327 การผลิตชิ้นส่วนสำหรับจักรยานยนต์และจักรยาน	1988	77.8	95.5
	1989	123.0	150.9
	1990	98.6	121.0
	1991	81.5	100.0

ขั้นตอนต่อไปคือ นำ SYPRO ที่ปรับฐานแล้วคูณกับค่าสัมประสิทธิ์การเปลี่ยนแปลง U_{jk} จากนั้นนำมารวมกันเพื่อให้ได้ WZ 93 class 35.42 โดยที่ปัจจัยการเปลี่ยนแปลง U_{jk} class ต่าง ๆ ของ WZ93 สามารถทำได้ดังนี้

ตาราง 5 การนำปัจจัยการเปลี่ยนแปลงไปใช้ (2) รวมดัชนี SYPRO ที่ปรับฐานแล้ว

กิจกรรมทางเศรษฐกิจ		ปัจจัยการเปลี่ยนแปลง U_{jk}	ดัชนีการผลิต (เดือน) 1991=100		
SYPRO	WZ 93		1988	1989	1990
3324		66.9	70.1	86.4	102.4
3327		33.1	95.5	150.9	121.0
	3542		78.5	107.0	108.6

ตัวอย่างที่ 2 การสำรวจการขายปลีกและการขายส่งรายเดือน (MWRTS) สถิติแคนาดา

เนื่องจากสถิติแคนาดามีการเปลี่ยนการจัดจำแนกประเภทอุตสาหกรรมจากการจัดจำแนกมาตรฐานอุตสาหกรรม (the Standard Industrial Classification : SIC) เป็นการจัดจำแนกประเภทอุตสาหกรรมในอเมริกาเหนือ (the North American Industry Classification System : NAICS) ทำให้โครงการการสำรวจการขายปลีกและขายส่ง (Monthly Wholesale and Retail Survey : MWRTS) ของสถิติแคนาดา ที่พัฒนาขึ้นในช่วงปลายทศวรรษที่ 1980 เพื่อผลิตค่าประมาณการขายสินค้าและคลังสินค้าตามภาคอุตสาหกรรมที่จัดจำแนกตาม SIC ต้องเปลี่ยนการจัดจำแนกเป็นจัดจำแนกตาม NAICS แต่เนื่องจากระบบการสำรวจที่มีอยู่ไม่ยินยอมให้หน่วยตัวอย่างจัดจำแนกตาม NAICS จึงต้องออกแบบการสำรวจให้ MWRTS ใหม่ โดยวางแผนแปลงและพยากรณ์ย้อนหลังใน ค.ศ. 2003 และดำเนินการเผยแพร่ข้อมูลตาม NAICS ที่อ้างอิงจากค่าประมาณ ในปลาย ค.ศ. 2003 เนื่องจากการแบ่งชั้นและการสุ่มตัวอย่างของ MWRTS มีการปรับปรุงในปี 1998 ดังนั้นจึงมี 2 ส่วนถูกนำไปใช้ ส่วนแรกคือช่วงก่อน ค.ศ. 1998 (เพื่อให้ส่วนนี้สอดคล้องกับผลลัพธ์ใน ค.ศ. 1998) และอีกส่วนเริ่มจาก ค.ศ. 1998 เป็นต้นไป โดยใช้วิธีมหภาคในการดำเนินงาน

การประมาณค่าของสัมประสิทธิ์การเปลี่ยนแปลง

ค่าสัมประสิทธิ์การเปลี่ยนแปลง $\alpha_{ij}(a,m)$ แสดงร้อยละของ group i ทั้งหมด (การจัดจำแนกเดิม) ที่ถูกจัดให้อยู่ group j (การจัดจำแนกใหม่) ในปีที่ a และเดือนที่ m

ค่าสัมประสิทธิ์การเปลี่ยนแปลง สำหรับ MWRTS ได้มาจากการเก็บตัวอย่าง 48 เดือน (4 ปี) ในช่วง มกราคม 1998 ถึง ธันวาคม 2001 โดยผู้เชี่ยวชาญจะพิจารณาค่าสัมประสิทธิ์ที่ใช้ไม่ได้หรือต่ำกว่า 0.3 % ในค่าสัมบูรณ์ โดยจะตัดสัมประสิทธิ์นั้นออกและทำการจัดสรรใหม่ ทำให้จำนวนหน่วยในอนุกรมลดลงจาก 1000 เป็น 230 จากนั้นอนุกรมที่เหลือจะถูกวิเคราะห์ด้วยกราฟิกเพื่อตรวจสอบว่ามีความแตกต่างในภูมิภาค ฤดูกาล หรือค่าผิดปกติหรือไม่ และสุดท้ายได้ประมาณค่าสัมประสิทธิ์เป็นรายเดือน เพื่อนำสัมประสิทธิ์ที่ได้ไปใช้กับ ค.ศ. 1991 – 1997 โดยคำนวณจากค่าเฉลี่ยของสัมประสิทธิ์ในเดือนเดียวกันของ ค.ศ. 1998 – 2001 ในแต่ละภูมิภาค (แบ่งตามภูมิภาค)

สามารถคำนวณค่าสัมประสิทธิ์การเปลี่ยนแปลง ได้ดังสมการ

$$\hat{\alpha}_{ij}^{r}(1991,m)=\ldots=\hat{\alpha}_{ij}^{r}(1997,m)=\frac{1}{k}\sum_{a=1998}^{2001}\delta_{ij}^{r}(a,m)\alpha_{ij}^{r}(a,m) \text{ โดยที่ } m=1,\ldots,12$$

โดยที่ k มีค่าเท่ากับ ผลรวม (ใน 4 ปี) ของตัวแปร $\delta_{ij}^{r}(a,m)$

$$\delta_{ij}^{r}(a,m)=\begin{cases}0 & \text{ถ้า } \alpha_{ij}^{r}(a,m) \text{ ไม่ถูกต้อง}\\ 1 & \text{อื่น ๆ}\end{cases}$$

สมการข้างต้นจะช่วยดึงค่าผิดปกติออกจากการคำนวณค่าเฉลี่ย ซึ่งสัมประสิทธิ์ที่ได้จะถูกปรับให้เป็น 100 % ในแต่ละปีที่ a เดือนที่ m ภูมิภาคที่ r และกลุ่มของการขายที่ i ตามการจัดจำแนกเดิม (SIC)

การประมาณค่าการจัดจำแนกเดิมด้วยสัมประสิทธิ์การเปลี่ยนแปลง

ในการประมาณค่ากลุ่มการขายให้เป็นการจัดจำแนกใหม่จะใช้ทฤษฎีสมการเชิงเส้นถ่วงน้ำหนัก (WLC) กับยอดรวมของแต่ละกลุ่มของการขายตามการจัดจำแนกเดิม โดยที่ค่าประมาณแต่ละกลุ่มการขายตามการจัดจำแนกใหม่ สามารถคำนวณได้จากสมการ

$$X_j(a,m)=\sum_i \alpha_{ij}(a,m)X_i(a,m)$$

โดยที่

$X_j(a,m)$ คือ กลุ่มการขาย j ตามการจัดจำแนกใหม่สำหรับปีที่ a และ เดือนที่ m

$X_i(a,m)$ คือ ยอดรวมของกลุ่มการขาย i ตามการจัดจำแนกเดิม

$\alpha_{ij}(a,m)$ คือ สัมประสิทธิ์การเปลี่ยนแปลง

ความต่อเนื่องของอนุกรมภายใต้ NAICS

อนุกรมตามการจัดจำแนกใหม่ (NAICS) แบ่งออกเป็น 3 ส่วน ส่วนแรกตั้งแต่เดือนมกราคม ค.ศ. 1991 ถึง เดือนธันวาคม ค.ศ. 1997 ซึ่งค่าประมาณในช่วงนี้คำนวณได้จากการถ่วงน้ำหนักด้วยค่าประมาณสัมประสิทธิ์การเปลี่ยนแปลง ส่วนที่สองจะเริ่มจากเดือนมกราคม ค.ศ. 1998 ถึงเดือนที่ยกเลิกการสำรวจด้วยกลุ่มตัวอย่างเดิม (สำรวจแบบเก่า) ในส่วนที่สองอนุกรมตาม NAICS จะได้ค่าประมาณจากตัวอย่างที่ถูกเก็บโดยตรงหรือค่าสังเกตของสัมประสิทธิ์การเปลี่ยนแปลง (เนื่องจากเป็นช่วงเวลาที่ลงรหัสคู่)

ส่วนที่สามเริ่มต้นตั้งแต่การสำรวจใหม่ โดยได้ทำการสำรวจซ้อนทับกัน 2-3 เดือน ทั้งการสำรวจแบบเดิมและแบบใหม่ (ทดสอบการเก็บแบบขนาน) เนื่องจากคาดว่าในช่วงเวลานี้อาจเกิดการแบ่งช่วงของอนุกรมจากการเปลี่ยนไปใช้แบบสำรวจใหม่ได้ และเนื่องจากการแบ่งช่วงไม่ได้เกิดจากการเปลี่ยนแปลงการจัดจำแนกเท่านั้น แต่อาจเกิดจากการที่มีการเปลี่ยนแปลงอื่น ๆ ในระเบียบวิธีด้วย ซึ่งเมื่อพิจารณาแล้วเห็นว่าควรนำข้อมูลในช่วงที่ทดสอบแบบมาใช้กับอนุกรมที่พยากรณ์ย้อนหลังด้วย โดยใช้ constant multiplicative ปรับตลอดช่วงเวลาเพื่อปรับปรุงข้อมูลในอดีตให้เป็นไปตามการสำรวจใหม่

นอกจากนี้ยังมีการแบ่งช่วงในเดือนมกราคม ค.ศ. 1998 เนื่องจากเปลี่ยนจากการใช้ค่าจาก ค่าประมาณสัมประสิทธิ์การเปลี่ยนแปลง เป็นการใช้ค่าสังเกตที่สำรวจได้โดยตรง ซึ่งในการลดผลกระทบนี้ จะพยากรณ์ย้อนหลังให้ ค.ศ. 1998 ด้วย โดยใช้ค่าประมาณสัมประสิทธิ์การเปลี่ยนแปลง (เช่น ค่าสัมประสิทธิ์ที่คำนวณจากค่าเฉลี่ยของ 4 ปี ซึ่งรวม ค.ศ. 1998 ด้วย) ซึ่งในความเป็นจริงค่าสัมประสิทธิ์ใน ค.ศ. 1998 มีความแตกต่างจากค่าเฉลี่ยของ 3 ปีที่เหลือ โดยมีความเสถียรมากกว่า เนื่องจากในการคำนวณค่าเฉลี่ยจะมีการลบค่าผิดปกติออกไป และการขยายการประมาณค่าส่วนแรกของอนุกรม (ตั้งแต่ มกราคม 1991- ปีที่ยกเลิกการสำรวจแบบเก่า) ออกมาจนถึงธันวาคม ค.ศ. 1998 จะทำให้การแบ่งช่วงระหว่างส่วนแรกและส่วนที่สองหายไป

แหล่งที่มาของข้อผิดพลาด

ข้อผิดพลาดประการแรกที่อาจเกิดขึ้นคือกรอบตัวอย่างของข้อมูล อาจมีการจัดจำแนกที่ผิดพลาดในช่วงระหว่าง ค.ศ. 1998 ถึง ค.ศ. 2001 ซึ่งไม่เพียงแต่ส่งผลกระทบต่อตัวมันเองเท่านั้น แต่ยังส่งผลต่อข้อมูลในเดือนเดียวกันในช่วงระหว่าง ค.ศ. 1991 ถึง ค.ศ. 1997 (ถูกพยากรณ์ย้อนหลัง) อีกด้วย เพื่อลดผลกระทบจากหน่วยสถิติที่มีการจัดจำแนกผิด ควรมีการตรวจสอบและลงรหัสใหม่ให้กับสถานประกอบการรายใหญ่ เพื่อลดผลกระทบดังกล่าว

ข้อผิดพลาดประการที่สองมาจากการใช้ค่าสัมประสิทธิ์การเปลี่ยนแปลง ที่คำนวณในช่วงปีปัจจุบัน (ค.ศ. 1998 - 2001) ไปประมาณค่าให้กับปีในอดีตที่มีระยะเวลานาน วิธีการนี้จะเหมาะสมเมื่อการกระจายตัวของอนุกรมตามการจัดจำแนกเดิมจากปีหนึ่งไปยังอีกปีหนึ่งคงที่ แต่ถ้าไม่คงที่เรายังสันนิษฐานได้ว่าความเสี่ยงของข้อผิดพลาดใน ค.ศ. 1997 จะน้อยกว่า ค.ศ. 1991 ซึ่งโดยส่วนใหญ่สมมติฐานความคงที่ได้รับการยอมรับ

อย่างไรก็ตามการใช้สัมประสิทธิ์การเปลี่ยนแปลง ดังที่อธิบายไว้ข้างต้นนั้นอาจไม่สามารถใช้ได้ (หรือไม่ควรใช้) ในกรณีที่อุตสาหกรรมมีการเปลี่ยนแปลงที่สำคัญ เช่น กรณีที่อุตสาหกรรมหนึ่งไม่ได้มีความสำคัญมากนักในช่วงเวลาเริ่มต้นที่เก็บข้อมูลแต่กลับมีความสำคัญในปัจจุบัน ดังนั้นค่าสัมประสิทธิ์การเปลี่ยนแปลง ที่จะนำไปประมาณค่าในปีเริ่มต้นจะต้องมีการปรับให้ลดลง ในขณะที่ค่าสัมประสิทธิ์ของ sector อื่น ๆ ต้องได้รับการปรับให้สูงขึ้น ซึ่งผู้เชี่ยวชาญจะเป็นผู้พิจารณาค่าที่ปรับ และอ้างอิงข้อมูลระดับจุลภาคตามการจัดจำแนกใหม่บางส่วน ซึ่งการปรับปรุงประเภทนี้สามารถสร้างเป็นแบบจำลองรูปแบบต่าง ๆ ตามช่วงเวลาของสัมประสิทธิ์ได้

การเกิดข้อผิดพลาดอีกประการคือ การใช้ค่าสัมประสิทธิ์ที่อ้างอิงจากตัวแปรหนึ่งไปใช้กับตัวแปรอื่น ซึ่งในกรณี MWRTS มีตัวแปรที่น่าสนใจสองตัวแปรคือ ยอดขาย และคลังสินค้า ในกรณีนี้จะคำนวณค่าสัมประสิทธิ์ทั้งหมดโดยอ้างอิงจากยอดขาย โดยที่อนุกรมของคลังสินค้านั้นใช้สัดส่วนของอนุกรมของยอดขายเช่นกัน

ตัวอย่างที่ 3 การสำรวจเงินเดือนและชั่วโมงการทำงาน (SEPH) - สถิติแคนาดา

ในตัวอย่างนี้อธิบายการแปลงชุดข้อมูลในอดีตจากการจัดประเภทมาตรฐานอุตสาหกรรม 1980 (SIC80) เป็นการจัดจำแนกประเภทอุตสาหกรรมในอเมริกาเหนือ (NAICS) โดยใช้ข้อมูลจาก 2 ช่วงเวลาคือ ข้อมูล 3 เดือนที่ต่อกันใน ค.ศ. 1998 และข้อมูล 3 เดือนที่ต่อกัน ใน ค.ศ. 1999 โดยใช้ข้อมูลทะเบียนจาก Statistics Canada Business Register เพื่อลงรหัสใหม่ในระดับข้อมูลจุลภาค (ระดับสถานประกอบการ) จาก SIC80 เป็น NAICS จากนั้นคำนวณค่าประมาณเพื่อรวมแต่ละกิจกรรมของอุตสาหกรรมตาม SIC ให้เป็นกิจกรรมของอุตสาหกรรมตาม NAICS โดยแยกแต่ละจังหวัดและแต่ละตัวแปร[10] โดยที่ค่าประมาณเหล่านี้จะนำไปใช้เป็นสัมประสิทธิ์การเปลี่ยนแปลงที่อยู่ในระดับเดียวกันกับ SIC80 (ตามจังหวัดและตัวแปร) สัดส่วนเหล่านี้ใช้เพื่อแปลงค่าประมาณ SIC80 เป็นค่าประมาณ NAICS ตลอดช่วงเวลาตั้งแต่ มกราคม 1991 ถึง ธันวาคม 2000

เมื่อแปลงข้อมูลแล้ว จะนำอนุกรมใหม่มาวิเคราะห์ความสอดคล้อง รวมถึงนำการแก้ไขข้อมูลในอดีต (ตั้งแต่เริ่มต้นการออกแบบใหม่ในเดือนพฤษภาคม 1996) มารวมในอนุกรมข้อมูลด้วย และเนื่องจากจำนวนของอนุกรมนั้นเกี่ยวข้องกับตัวแปรมากมาย ดังนั้นจึงมุ่งเน้นการวิเคราะห์ที่ตัวแปรที่มีนัยสำคัญที่สุดสำหรับแต่ละจังหวัด (เช่น รายได้เฉลี่ยรายสัปดาห์ การจ้างงาน ฯลฯ)

เจ้าของโครงการควรหมายเหตุว่าวิธีการแปลงที่ใช้มีข้อจำกัดบางประการ เช่น การไม่มีข้อมูลสถานประกอบการตาม NAICS (ไม่พร้อมใช้) สำหรับสถานประกอบการในช่วงเวลาก่อนหน้า รวมถึงการเปลี่ยนแปลงของประชากรเป้าหมาย (สถานประกอบการกับลูกจ้าง) ในช่วงเวลาที่ผ่านมา (การเพิ่มและการลดของสถานประกอบการ) หากมีข้อมูลที่ลงรหัสใหม่แต่ละเดือน อาจให้ผลลัพธ์ที่เห็นว่ามีอะไรเปลี่ยนแปลง ตัวอย่างเช่น มีอุตสาหกรรมใหม่ ๆ ที่เกิดขึ้นภายในรอบทศวรรษ ในขณะที่บางอุตสาหกรรมอาจหายไปในบางจังหวัด ซึ่งในส่วนนี้มีผลกระทบเชิงลบต่อคุณภาพของอนุกรมที่แปลง โดยเฉพาะอย่างยิ่ง ช่วงเวลาที่ถูกประมาณค่าที่ห่างออกไป

นอกจากนี้ จากการใช้วิธีสัดส่วนการเปลี่ยนแปลงอาจทำให้เกิดรูปแบบอนุกรมมีความสัมพันธ์กันมากใน NAICS 4 หลัก โดยจะใกล้เคียงกันมากในช่วง ค.ศ. 1991 - 2000 เนื่องจากถูกแปลงตามข้อมูลรวมยอด (กลุ่มใหญ่) ในกรณีนี้ความสัมพันธ์ที่ใกล้เคียงกันอาจสิ้นสุดลงในเดือนมกราคม ค.ศ. 2001 เนื่องจากมีการวิเคราะห์แยกกันในแต่ละหมู่ในอนุกรม

[10] SEPH สร้างค่าประมาณที่อ้างอิงจากตัวแปร 11 ค่า จากการคำนวณตัวแปรอื่น ๆ ทั้งหมด