การคาดการณ์ประชากร (2):
แบบจำลองเชิงเปรียบเทียบ แบบจำลองเชิงอัตราส่วน
และการเลือกแบบจำลองที่เหมาะสมในการคาดการณ์

โดย รศ. ดร. วรรณศิลป์ พีรพันธุ์ © 2001-2003

แบบจำลองเชิงเปรียบเทียบ (Comparative Model)

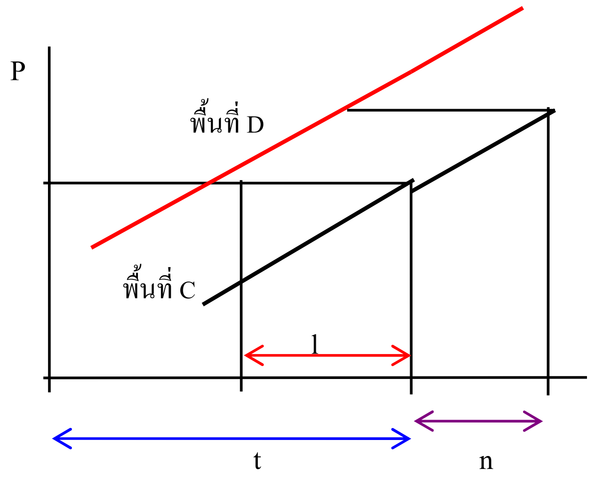
แบบจำลองเชิงเปรียบเทียบตั้งอยู่บนสมมุติฐานว่า การเติบโตของประชากรในอนาคตของพื้นที่หนึ่ง ซึ่งในที่นี้จะขอเรียกว่าพื้นที่ C สามารถคาดการณ์ได้โดยอาศัยพื้นฐานความรู้เกี่ยวกับการเติบโตของประชากรในอดีตของอีกพื้นที่หนึ่ง ซึ่งพื้นที่หลังนี้จะขอเรียกว่าพื้นที่ D หรือเรียกโดยทั่วไปว่า “Pattern Area” โดยพื้นที่ D และ C ต่างมีต้นกำเนิดที่เหมือนกัน และในอดีตที่ผ่านมาก็มีรูปแบบการเติบโตของประชากรในลักษณะเหมือน ๆ กันด้วย แต่เหตุการณ์บน D เกิด ขึ้นก่อน C เราจึงสามารถใช้ D เป็นตัวเทียบเคียงในการคาดการณ์ประชากรในอนาคตของ C ได้
แนวความคิดนี้มักไม่ค่อยจะมีการแสดงออกทางคณิตศาสตร์บ่อยครั้งนัก แต่ก็สามารถเขียนเป็นสมการได้ดังนี้:

$$ {}_{C}P_{t+n} = (p)\ {}_{D}P_{(t-l)+n} \qquad \text{..............................} \ [1] $$

เมื่อ: ${}_{C}P$ = ประชากรของพื้นที่ C
${}_{D}P$ = ประชากรของพื้นที่ D
t = ช่วงเวลาปัจจุบัน
l = ช่วงห่างของการพัฒนาระหว่างสองพื้นที่
n = ช่วงเวลาต่อจากปัจจุบันที่ต้องการคาดการณ์ประชากรของพื้นที่ C
p = สัดส่วนของ ${}_{C}P/{}_{D}P$

ตราบใดที่เรามีเหตุผลที่เชื่อถือได้ว่าพื้นที่ที่เราศึกษา และ Pattern Area มีความต่อเนื่องของอิทธิพลที่เปลี่ยนไปเหมือน ๆ กัน โดยที่พื้นที่ที่เรากำลังศึกษากำลังเดินตามหลังอยู่ ทฤษฎีของการเทียบเคียงก็ยังคงน่าเชื่อถือ และการคาดการณ์แบบเปรียบเทียบก็ยังคงใช้ได้ ในอดีตนั้นเทคนิคนี้ถูกนำมาประยุกต์ใช้บ่อย ๆ กับเมืองทั้งเมือง โดยมีสมมุติฐานว่า เมืองที่ศึกษาจะมีรูปแบบการ เติบโตของประชากรตามรอยของอีกเมืองหนึ่ง ซึ่งการนำเทคนิคดังกล่าวมาใช้ในลักษณะเช่นนี้มักจะไม่ค่อยประสบความสำเร็จเท่าใดนัก แต่ถ้าหากเรานำมาประยุกต์ใช้เฉพาะกับพื้นที่ส่วนย่อยภายในเมืองแล้ว การคาดการณ์อาจมีความถูกต้องได้ในระดับสูงทีเดียว ตัวอย่างเช่น บริเวณชาน เมือง

ตรงรอยต่อระหว่างเมืองกับชนบท มักมีแนวโน้มของการเติบโตตามแบบของพื้นที่ที่อยู่ถัด เข้ามาในเมือง เป็นต้น



แผนภูมิแสดงตัวอย่างของการใช้แบบจำลองเชิงเปรียบเทียบ ค่าสัดส่วน (p) = 1

แบบจำลองเชิงอัตราส่วน (Ratio Model)

แนวความคิดพื้นฐานของแบบจำลองเชิงอัตราส่วน สามารถอธิบายได้โดยตัวอย่างต่อ ไปนี้ สมมุติว่าการเติบโตของประชากรในพื้นที่ที่เราทำการศึกษา ซึ่งในที่นี้จะขอเรียกว่าชุมชน C แปรผันไปตามการเติบโตของประชากรในภูมิภาคที่ชุมชนนี้ตั้งอยู่ ซึ่งในที่นี้จะขอเรียกว่าภูมิภาค S ถ้าเรามั่นใจในสมมุติฐานที่ว่า ลักษณะความสัมพันธ์เช่นนี้จะยังคงดำเนินต่อไปในอนาคตอัน ใกล้ และขณะเดียวกันเรามีข้อมูลการคาดการณ์ประชากรของภูมิภาค S อยู่ในมือแล้ว เราก็สามารถจะคำนวณส่วนแบ่งของประชากรในชุมชนที่ศึกษาออกจากจำนวนประชากรของภูมิภาคได้

สมมุติฐานดังกล่าวก่อให้เกิดรูปแบบของสมการได้ดังต่อไปนี้:

$$\frac{{}_{C}P_{t+n}}{{}_{S}P_{t+n}} = \frac{{}_{C}P_{t}}{{}_{S}P_{t}}$$

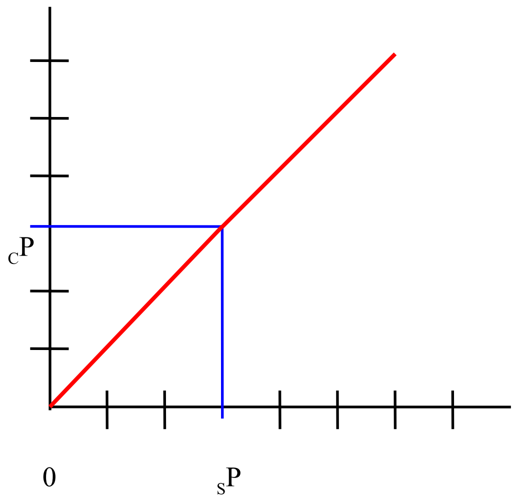
และเพื่อใช้ในการคาดการณ์ได้ง่ายขึ้น เราสามารถจัดรูปสมการดังกล่าวเสียใหม่โดยการคูณ ${}_{S}P_{t+n}$ เข้าทั้งสองข้าง ซึ่งจะได้รูปแบบดังนี้:

$$ {}_{C}P_{t+n} = \left[\frac{{}_{C}P_{t}}{{}_{S}P_{t}} \right] {}_{S}P_{t+n} \quad \text{..........................} \quad [2] $$

เมื่อพิจารณาสมการดังกล่าวจะพบว่ามีลักษณะคล้ายกับแบบจำลองเชิงเส้นตรง (Linear Model) ทั้งนี้เนื่องจากในความเป็นจริงแล้วมันก็คือแบบจำลองเชิงเส้นตรงนั่นเอง โดยจากสมการ $Y_c = a + bX$ เมื่อ a หรือค่าตัวคั่น (Intercept) ถูกสมมุติให้มีค่าเป็นศูนย์และตัวแปรอิสระ (X) ได้แก่ จำนวนประชากรของภูมิภาค ขณะที่ค่าสัมประสิทธิ์ b ได้แก่ อัตราส่วนของประชากร (${}_{C}P_{t}/{}_{S}P_{t}$)

ข้อมูลในอดีตที่จำเป็นต้องใช้มีเพียงจำนวนประชากรของชุมชนและของภูมิภาคในช่วง เวลาใดช่วง เวลาหนึ่ง เมื่อพล็อตจุดดังกล่าวลงในกราฟแล้วลากเส้นเชื่อมจากจุดนั้นไปยังจุดกำเนิด (Origin) 0,0

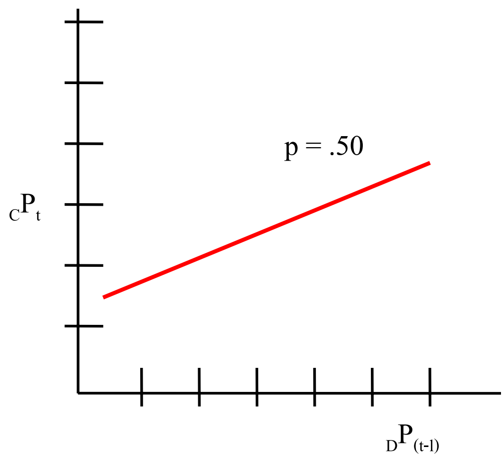
เราก็สามารถต่อแนวเส้นนั้นออกไปในอนาคตได้



แผนภูมิแสดงความสัมพันธ์ระหว่างชุมชนที่ศึกษากับภูมิภาคที่ชุมชนตั้งอยู่

แม้ว่าโดยพื้นฐานทางสถิติแล้ว แบบจำลองเชิงอัตราส่วน จะดูด้อยกว่าแบบจำลอง เชิง เส้นตรงซึ่งมีพื้นฐานของข้อมูลที่มีรายละเอียดมากกว่า แต่เราก็สามารถพัฒนาให้แบบจำลองเชิงอัตราส่วนมีความถูกต้องยิ่งขึ้นได้ โดยพิจารณาข้อมูลจากหลาย ๆ ช่วงเวลา และใช้ค่าเฉลี่ย หรือแนวโน้มของอัตราส่วนเหล่านั้นแทน ซึ่งก็หมายความว่าเรากำลังใช้วิธีทางคณิตศาสตร์ในทำนองเดียวกับที่เราใช้ในแบบจำลองเชิงเส้นตรงนั่นเอง

ความแตกต่างเพียงอย่างเดียวระหว่างแบบจำลองเชิงอัตราส่วน และแบบจำลองเชิงเส้นตรงก็คือ ตัวแปรต้นเหตุที่แตกต่างกัน ทำให้แบบจำลองเชิงอัตราส่วนมีความได้เปรียบในแง่ของความเชื่อมั่นในการคาดการณ์อนาคตของชุมชน ซึ่งมักจะถูกกำหนดโดยระบบที่มีขนาดใหญ่ กว่า และอิทธิพลจากปัจจัยทางสังคม เศรษฐกิจ การเมือง และกายภาพของระบบนั้น



แผนภูมิแสดงแบบจำลองเชิงเปรียบเทียบ เมื่อนำมาดัดแปลงให้อยู่ในรูปของแบบจำลองเชิงอัตราส่วน

การที่นักวางแผนสามารถใช้ประโยชน์จากการคาดการณ์ในระดับใหญ่ ซึ่งเชื่อถือได้ค่อนข้างแน่นอนมาเป็นพื้นฐานสำหรับการคาดการณ์ในระดับย่อยของพื้นที่ศึกษา ทำให้แบบจำลองเชิงอัตราส่วนเป็นที่นิยมใช้กันมากกว่าแบบจำลองเชิงเปรียบเทียบ ซึ่งมีลักษณะสมการทางพีชคณิตเกือบจะเหมือนกันทุกอย่าง กล่าวคือ ตัวแปรอิสระ ได้แก่ จำนวนประชากรของพื้นที่ อื่น และช่วงห่างของการพัฒนา จำนวนประชากรของพื้นที่ศึกษาในแต่ละช่วงเวลาอยู่ในรูปของอัตราส่วนกับจำนวนประชากรของพื้นที่เทียบเคียง ส่วนความแตกต่างระหว่างแบบจำลองทั้งสองก็คือ แบบจำลองเชิงอัตราส่วนสามารถสลับบทบาทระหว่างพื้นที่ได้ โดยการเติบโตของประชากรในอนาคตของพื้นที่ที่เล็กกว่า เช่น กรุงเทพมหานคร อาจนำมาใช้คาดการณ์การเติบโต ของประชากรในอนาคตของภาคมหานครได้ หรือการเติบโตของประชากรในส่วนหนึ่งของเมือง ก็อาจจะนำมาใช้คาดการณ์การเติบโตของประชากรทั้งเมืองได้ นอกจากนี้แบบจำลองเชิงอัตราส่วนอาจนำมาใช้คาดการณ์ข้อมูลด้านอื่นโดยเทียบให้เป็นอัตราส่วนกับจำนวนประชากรของพื้นที่ได้ เช่น การเทียบ อัตราส่วนของโรงเรียนกับจำนวนประชากรในชุมชน หรือ การเทียบอัตราส่วนของการจ้างงานกับจำนวนประชากร เป็นต้น

การเลือกแบบจำลองที่เหมาะสมในการคาดการณ์ประชากร

ปัญหาสำคัญประการหนึ่งที่มักจะพบโดยทั่วไปในการคาดการณ์ประชากรก็คือ การเลือกแบบจำลองที่เหมาะสม วิธีการเบื้องต้นที่สามารถใช้ในการแก้ปัญหาก็คือ การพล็อตข้อมูลที่ เรา มีอยู่ลงในกราฟ และดูลักษณะเส้นกราฟนั้นว่าเข้ากับแบบจำลองใด ถ้าเรามีแต่เพียงข้อมูลในอดีตของชุมชนที่ศึกษาเพียงอย่างเดียว เราก็สามารถตัดแบบจำลองเชิงเปรียบเทียบและแบบจำลองเชิงอัตราส่วนออกไปได้ แต่ถ้าเรามีข้อมูลของภูมิภาคหรือพื้นที่เทียบเคียงอยู่ เราก็ควรจะ พิจารณาแบบจำลองทั้งสองด้วย นอกจากการพิจารณาความเหมาะสมของแบบจำลองแล้ว การยอมรับสมมุติฐานของสมการต่าง ๆ ที่ใช้คาดการณ์ก็เป็นสิ่งสำคัญอีกประการหนึ่งที่จะต้องได้รับการพิจารณาและตัดสินใจ โดยใช้พื้นฐานความรู้ในสาขาอื่น ๆ ที่เกี่ยวข้องมาช่วยพิจารณา

นอกจากการพิจารณาจากกราฟแล้ว เรายังสามารถประเมินความเหมาะสมของแบบจำลองได้โดย การเปรียบเทียบผลการคาดการณ์กับข้อมูลในอดีตว่ามีความแตกต่างกันมากน้อย เพียงใด “Output Evaluation” เป็นการประเมินความเหมาะสมของแบบจำลอง โดยการเปรียบ เทียบข้อมูลในอดีตกับตัวเลขที่ได้จากการคาดการณ์โดยแบบจำลองต่าง ๆ โดยอาศัยหลักการที่ว่าแบบจำลองที่สามารถอธิบายเหตุการณ์ในอดีตได้ดีที่สุดย่อมจะสามารถใช้ในการคาดการณ์อนาคตได้ดีด้วย วิธีการที่นิยมใช้ในการวัดความถูกต้องของการคาดการณ์ที่นิยมใช้กัน ได้แก่ การหาค่าความคลาดเคลื่อนเฉลี่ย (Mean Error) และความคลาดเคลื่อนร้อยละสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error)

Mean Error (ME) มีสูตรในการคำนวณดังนี้:

$$\mathbf{ME} = \frac{\sum(Y_0 - Y_C)}{N} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ [3]$$

โดยที่: Y_0 = ตัวเลขประชากรจริง (Observed Value)
Y_C = ตัวเลขประชากรที่ได้จากการคาดการณ์โดยแบบจำลอง (Calculated Value)
N = จำนวนข้อมูล (ช่วงเวลา) ในอดีต

เนื่องจาก Mean Error เป็นค่าเฉลี่ยของความแตกต่าง ระหว่างตัวเลขจริงกับตัวเลขที่ได้จากการคาดการณ์โดยไม่ได้ยกกำลัง และไม่ได้ใช้ค่าสัมบูรณ์ จึงเป็นไปได้ที่ค่าความแตกต่างที่เป็นบวกและลบ จะหักลบกันไปจนทำให้ค่า Mean Error เป็นศูนย์ Mean Error จึงไม่ใช่เกณฑ์ที่ดีนัก ในการวัดความคลาดเคลื่อนของการคาดการณ์แต่ละค่า แต่เป็นประโยชน์ในแง่ของการมองภาพรวม ว่ามีความคลาดเคลื่อนอย่างเป็นระบบ (Systematic Error) เช่น คาดการณ์สูงเกินไปทั้งหมด หรือต่ำเกินไปทั้งหมดหรือไม่ ค่า Mean Error ที่เป็นศูนย์จะช่วยให้ทราบว่าผลของการคาดการณ์ ไม่มีความคลาดเคลื่อนในลักษณะดังกล่าว แม้ว่าเราจะยังไม่ทราบชัดว่าความคลาดเคลื่อนของการ คาดการณ์ในแต่ละค่าจะมากหรือน้อยเพียงใด

สำหรับ Mean Absolute Percentage Error (MAPE) มีสูตรในการคำนวณคือ:

$$\text{MAPE} = \frac{\sum (| Y_0 - Y_C | / Y) * 100}{N} \quad \text{.........................} \; [4]$$

จะเห็นได้ว่า ค่าของความแตกต่าง $Y_0 - Y_C$ ถูกทำให้เป็นค่าสัมบูรณ์โดยไม่คิดเครื่องหมายเพื่อให้เราสามารถวัดความคลาดเคลื่อนของการคาดการณ์แต่ละค่าได้ ความคลาดเคลื่อนดัง กล่าวถูกทำให้เป็นร้อยละเพื่อให้สามารถเปรียบเทียบได้ถูกต้องยิ่งขึ้นเนื่องจากไม่มีหน่วย ทำให้ไม่ได้รับผลกระทบจากขนาดของตัวเลข MAPE จึงเหมาะสมที่จะใช้ในการพิจารณาความคลาด เคลื่อนของการคาดการณ์แต่ละค่า และการเปรียบเทียบความแม่นยำของการคาดการณ์หลาย ๆ ชุด ที่อาจใช้ช่วงเวลาของการคาดการณ์ที่ต่างกัน