

# Symmetric and Asymmetric Measures of Association

สุทธิวรรณ พิรศักดิ์โสภณ \*

ฉัตรศิริ ปิยะพิมลสิทธิ์ \*\*

ในการหาความสัมพันธ์ระหว่างตัวแปร 2 ตัว มีวิธีการหาหลายวิธี ดังที่เราเคยได้ทราบกันไปแล้ว เช่น point biserial , phi ( $\phi$ ) , Spearman Rho ( $\rho$ ) หรือวิธีสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน ( $r_{xy}$ ) ซึ่งแต่ละวิธีมีข้อตกลงของการใช้แตกต่างกันไป นอกจากวิธีดังกล่าวเหล่านี้แล้ว Goodman และ Kruskal (1994) ยังได้เสนอวิธีการความสัมพันธ์ระหว่างตัวแปร 2 ตัว โดยที่ตัวแปรทั้งสองอยู่ในรูปแบบตารางการณัจจร (Contingency Table) ซึ่งเรียกว่าการหาความสัมพันธ์แบบ Symmetric และ Asymmetric

## Symmetric Measure of Association

เป็นการวัดความสัมพันธ์ระหว่างตัวแปร 2 ตัว โดยที่ตัวแปรแต่ละตัวมีระดับการวัดเป็นนามบัญญัติ และตัวแปรแต่ละตัวอาจแบ่งเป็นกลุ่มเป็นพวกได้ตั้งแต่ 2 กลุ่มขึ้นไป เช่น การหาความสัมพันธ์ระหว่างระดับการศึกษา และรายได้ ซึ่งระดับการศึกษาอาจแบ่งเป็น สูง กลาง และต่ำ รายได้อาจแบ่งเป็น สูง และ ต่ำ หรืออาจหาความสัมพันธ์ระหว่างชนชั้นทางสังคมกับระดับกิจกรรมเกี่ยวกับการเลือกตั้ง ซึ่งชนชั้นทางสังคมอาจแบ่งเป็น ต่ำ กลาง และสูง ระดับกิจกรรมเกี่ยวกับการเลือกตั้งอาจแบ่งเป็น สูง กลาง และต่ำ เป็นต้น ความสัมพันธ์ระหว่างตัวแปรทั้ง 2 ในลักษณะเช่นนี้เรียกว่าสัมประสิทธิ์แลมด้า (lambda coefficient  $\lambda$ )

---

\* ผู้ช่วยศาสตราจารย์ประจำภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

\*\* นิสิตปริญญาเอก สาขาการทดสอบและวัดผลการศึกษา มหาวิทยาลัยศรีนครินทรวิโรฒ

สูตรที่ใช้ในการคำนวณมีดังนี้

$$\lambda = \frac{\sum_{j=1}^J n_{mj} + \sum_{i=1}^I n_{im} - n_{m+} - n_{+m}}{2n - n_{m+} - n_{+m}} \dots\dots\dots(1.1)$$

- เมื่อ  $n_{mj}$  แทน ความถี่ที่มากที่สุดในแต่ละคอลัมน์  $j$
- $n_{im}$  แทน ความถี่ที่มากที่สุดในแต่ละแถว  $i$
- $n_{m+}$  แทน ผลรวมของแถวที่มีค่ามากที่สุด
- $n_{+m}$  แทน ผลรวมของคอลัมน์ที่มีค่ามากที่สุด
- $n$  แทน จำนวนข้อมูลที่สังเกตได้

ตัวอย่าง 1 จงหาความสัมพันธ์ระหว่างระดับการศึกษาและการเลือกอาชีพ จากข้อมูลในตารางต่อไปนี้

การเลือกอาชีพ	ระดับการศึกษา					รวม
	ต่ำกว่ามัธยม	มัธยม	ประกาศนียบัตร	ป.ตรี	สูงกว่าป.ตรี	
กรรมกร/ชานา	347	128	84	37	5	601
ทักษะฝีมือทางช่าง	164	277	103	43	36	623
เสมียน/บัญชี	30	77	217	147	80	551
วิชาชีพ/การจัดการ	2	34	82	198	267	583
รวม	543	516	486	425	388	2,358

จากข้อมูลในตัวอย่าง 1 คำนวณหาความถี่ที่มากที่สุดของแต่ละแถว แต่ละคอลัมน์ และผลรวมทั้งหมดในแต่ละแถว แต่ละคอลัมน์ที่มีค่ามากที่สุด จะได้ดังนี้

แต่ละคอลัมน์ความถี่ที่มากที่สุดคือ 347 277 217 198 267

$$\therefore \sum_{j=1}^J n_{mj} = (347 + 277 + 217 + 198 + 267) = 1,306$$

แต่ละแถวความถี่ที่มากที่สุดคือ 347 277 217 267

$$\therefore \sum_{i=1}^I n_{im} = (347 + 277 + 217 + 267) = 1,108$$

ผลรวมทั้งหมดในแต่ละแถวที่มากที่สุดและผลรวมทั้งหมดในแต่ละคอลัมน์ที่มากที่สุด คือ

$$\begin{aligned} n_{m+} &= 623 \\ n_{+m} &= 543 \\ n &= 2,358 \end{aligned}$$

แทนค่าที่ได้ลงในสมการ (1.1)

$$\begin{aligned} \lambda &= \frac{1,306 + 1,108 - 623 - 543}{2(2,358) - 623 - 543} \\ &= .352 \end{aligned}$$

จากตัวอย่างได้ค่าสัมประสิทธิ์แลมด้า ( $\lambda$ ) เท่ากับ .352 นั่นคือระดับการศึกษาและการเลือกอาชีพมีความสัมพันธ์กันทางบวกในระดับปานกลาง

#### Asymmetric measure of association

สัมประสิทธิ์แลมด้า ( $\lambda$ ) เป็นค่าที่บอกให้ทราบว่าตัวแปรทั้ง 2 มีความสัมพันธ์กันในทิศทางใดและขนาดเท่าใด แต่จะไม่บอกให้ทราบว่าตัวแปรหนึ่งจะพยากรณ์หรือทำนายตัวแปรอีกตัวหนึ่งได้ ดังนั้นในลักษณะข้อมูลแบบเดียวกับการหาสัมประสิทธิ์แลมด้า หากเราต้องการทำนายตัวแปรหนึ่ง (Y) ด้วยตัวแปรอีกตัวหนึ่ง (X) ซึ่งข้อมูลมีลักษณะเป็นตารางการแจกแจงเราสามารถใช้ในการวัดความสัมพันธ์แบบ Asymmetric ( $\lambda_y$ ) หรือที่เรียกว่า ดัชนีความสัมพันธ์เชิงพยากรณ์ (Index of predictive association) ซึ่งมีสูตรที่ใช้ในการคำนวณดังนี้

$$\lambda_y = \frac{\sum_{j=1}^J n_{mj} - n_{m+}}{n - n_{m+}} \dots \dots \dots (1.2)$$

- เมื่อ  $n_{mj}$  แทน ความถี่ที่มากที่สุดในแต่ละคอลัมน์ j
- $n_{m+}$  แทน ผลรวมของแถวที่มากที่สุด
- $n$  แทน จำนวนข้อมูลที่สังเกตได้

จากตัวอย่าง 1 ถ้าให้ Y แทนการเลือกอาชีพ และ X แทน ระดับการศึกษา หากต้องการทำนายการเลือกอาชีพจากระดับการศึกษา แทนค่าสูตร (1.2) จะได้ดังนี้

$$\lambda_y = \frac{1,306 - 623}{2,358 - 623} = 0.394$$

ดังนั้นสรุปได้ว่า จากข้อมูลที่เก็บรวบรวมมาได้คนที่มีการศึกษาสูงขึ้นการเลือกอาชีพจะเป็นพวกทำงานในตำแหน่งเสมียน/บัญชี หรือวิชาชีพ/การจัดการ ในขณะที่พวกที่มีการศึกษาน้อยการเลือกอาชีพจะเป็นพวกกรรมกร/ชาวนา

หรืออาจกล่าวได้ว่าระดับการศึกษาสามารถทำนายการเลือกอาชีพได้ประมาณ 39%

ในทางกลับกันจากตัวอย่าง 1 หากเราต้องการทำนายระดับการศึกษาจากการเลือกอาชีพก็สามารถทำนายได้เช่นเดียวกัน โดยใช้สูตร

$$\lambda_x = \frac{\sum_{i=1}^I n_{im} - n_{+m}}{n - n_{+m}} \dots\dots\dots(1.3)$$

จากตัวอย่าง 1 สามารถคำนวณหา  $\lambda_x$  ได้โดยแทนค่าลงในสูตร (1.3)

$$\lambda_x = \frac{1,108 - 543}{2,358 - 543} = 0.311$$

ดังนั้นสามารถทำนายระดับการศึกษาจากการเลือกอาชีพได้ประมาณ 31%

สัมประสิทธิ์  $\lambda_y$  มีขนาดระหว่าง 0 ถึง 1 ทั้งนี้ก็ขึ้นอยู่กับขนาดหรือผลรวมในแนวแถวหรือแนวคอลัมน์  $\lambda_y$  จะมีค่าเป็น 1 ในกรณีที่ถูก ๆ ค่าของกลุ่มย่อยในตัวแปรอิสระตกอยู่ในแต่ละกลุ่มย่อยของตัวแปรตาม และ  $\lambda_y$  จะมีค่าเป็น 0 ในกรณีที่แต่ละกลุ่มย่อยที่มีความถี่มากที่สุดของตัวแปรอิสระตกอยู่ในกลุ่มย่อยเดียวกันของตัวแปรตาม ดังข้อมูลในตัวอย่าง 2 ค่า  $\lambda_y$  เป็น 0

## ตัวอย่าง 2 ตารางแสดงชนิดของความผิดปกติกับอายุของผู้ป่วย

ชนิดของความ พิการ (Y)	อายุ (X)			รวม
	ต่ำกว่า 15 ปี	15-24 ปี	ตั้งแต่ 25 ปี	
ทางตา	327	497	4949	5773
ทางการได้ยิน	468	462	7944	8874
ทางการพูด	546	134	520	1200
ทางกระดูก	1258	2944	19145	23347
รวม	2599	4037	32558	39194

จากข้อมูลในตัวอย่าง 2 นี้ ผลรวมของความถี่ที่มากที่สุดในแต่ละคอลัมน์คือ

$$\therefore \sum_{j=1}^J n_{mj} = (1,258 + 2,944 + 19,145) = 23,347$$

ผลรวมของแถวที่มีความถี่มากที่สุดคือ

$$\therefore n_{m+} = 23,347$$

เมื่อแทนค่าลงในสูตร (1.2) จะทำให้ค่า  $\lambda_{Y}$  เป็น 0 นั้นแสดงว่า อายุไม่สามารถทำนายชนิดของความพิการได้

## สรุป

การหาค่าสัมประสิทธิ์แลมด้าเป็นการหาความสัมพันธ์ระหว่างตัวแปรสองตัวโดยที่ตัวแปรแต่ละตัวแบ่งเป็นกลุ่มหรือพวกย่อย ลักษณะข้อมูลเป็นแบบความถี่นั้นคือมีระดับการวัดเป็นแบบนามบัญญัติ และข้อมูลอยู่ในรูปตารางการแจกแจง นอกจากนี้แล้วเรายังสามารถนำไปหาค่าดัชนีความสัมพันธ์เชิงพยากรณ์สำหรับตัวแปรประเภทจัดกลุ่มได้อีกด้วย ในทางการวัดผลการศึกษารวมเราสามารถนำไปประยุกต์ใช้เพื่อหาดัชนีความสอดคล้องระหว่างผู้ประเมินได้อีกด้วย

## หนังสืออ้างอิง

- Guilford J.P. and Benjamin Fruchter. (1978). Fundamental Statistics in Psychology and Education. McGraw-Hill.
- Hays, L William. (1993). Statistics. 5<sup>th</sup> ed. Holt, Rinehart and Winston.
- Wiersma, Hinkle and Jurs. (1994). Applied Statistics for the Behavioral Sciences. 3<sup>rd</sup> ed. Houghton Mifflin.
- Loether and McTatish. (1993). Descriptive and Inferential Statistics: An Introduction. 4<sup>th</sup> ed. Allyn and Bacon.