

การแปลความหมายระดับความสัมพันธ์จากการใช้สถิติวิเคราะห์

บุญชม ศรีสะอาด¹

การศึกษาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ เป็นเรื่องที่มีผู้สนใจทำการค้นคว้ามา ยาวนานและต่อเนื่อง สถิติที่ใช้เป็นเครื่องมือในการศึกษาความสัมพันธ์ ได้แก่ สหสัมพันธ์อย่างง่าย (Simple Correlation) ใช้ตัวย่อ r หรือ r_{xy} สหสัมพันธ์พหุคูณ (Multiple Correlation) ใช้ตัวย่อ R และสหสัมพันธ์คาโนนิคอล (Canonical Correlation) ใช้ตัวย่อ R_c สหสัมพันธ์อย่างง่ายใช้วิเคราะห์หาความสัมพันธ์ระหว่างตัวแปร 2 ตัว นิยมเขียนแทนด้วย X กับ Y สหสัมพันธ์พหุคูณใช้วิเคราะห์หาความสัมพันธ์ระหว่างกลุ่มตัวแปรทางซ้ายมือ ซึ่งมีตั้งแต่ 2 ตัวแปรขึ้นไป ($X_1, X_2, \dots$) กับตัวแปรทางขวามือที่มีเพียงตัวเดียว (Y) ส่วนสหสัมพันธ์คาโนนิคอลใช้วิเคราะห์หาความสัมพันธ์ระหว่างกลุ่มตัวแปรทางซ้ายมือ ซึ่งมีตั้งแต่ 2 ตัวขึ้นไป ($X_1, X_2, \dots$) กับกลุ่มตัวแปรทางขวามือที่มีตั้งแต่ 2 ตัวขึ้นไป ($Y_1, Y_2, \dots$) ในที่นี้จะกล่าวเฉพาะกรณีสหสัมพันธ์อย่างง่าย

การที่ตัวแปรสองตัวมีความสัมพันธ์กัน หมายถึงการที่ค่าวัดในตัวแปรทั้งสอง แปรผันด้วยกันซึ่งอาจเป็นไปในทางเดียวกัน (แปรผันตามกัน) เช่น ถ้าค่า X สูง ค่า Y ก็สูงด้วย ค่า X ต่ำ ค่า Y ก็ต่ำด้วย หรือเป็นไปในทางตรงกันข้ามกัน (แปรผันสวนกัน กลับกัน หรือผกผัน) เช่น ถ้าค่า X สูง ค่า Y จะต่ำ แต่ถ้าค่า X ต่ำ ค่า Y จะสูง เป็นต้น

ค่าสหสัมพันธ์อย่างง่าย (r) มีตั้งแต่ 0 ถึง 1 กรณีที่แปรผันตามกัน ค่า r จะเป็นค่าบวก ถ้าแปรผกผันสวนกันค่า r จะเป็นลบ ถ้า r เป็น 0 แสดงว่าไม่ได้แปรผันด้วยกัน นั่นคือไม่มีความสัมพันธ์กัน

เมื่อวิเคราะห์หาค่า r จากข้อมูล ค่า X และ Y ของกลุ่มตัวอย่างที่เก็บรวบรวมมาทุกคน ค่า r ที่ได้ชี้ถึงว่าตัวแปร X กับ Y มีความสัมพันธ์กันหรือไม่ ถ้าสัมพันธ์กัน สัมพันธ์ในทิศทางใด (สัมพันธ์ทางบวก หรือทางลบ) และในระดับใด ซึ่งมีวิธีการ พิจารณาความสัมพันธ์ และระดับความสัมพันธ์ ดังนี้

พิจารณาว่าค่า r นั้น มีนัยสำคัญ (Significance) หรือไม่ โดยนำค่า r ที่ได้จากการวิเคราะห์ไป เปรียบเทียบกับค่าวิกฤต (Critical Value) ถ้ามีค่าน้อยกว่าค่าวิกฤต แสดงว่าค่า r นั้น ไม่มีนัยสำคัญ ชี้ถึงว่าตัวแปรทั้งสองไม่สัมพันธ์กัน (ณ ระดับนัยสำคัญที่กำหนดไว้นั้น)

¹ รองศาสตราจารย์ ดร. คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

ตัวอย่างเช่น ค่า r จากการวิเคราะห์เท่ากับ .43 ค่าวิกฤตเท่ากับ .50 ค่า r จากการวิเคราะห์น้อยกว่าค่าวิกฤต จึงสรุปได้ว่า ตัวแปรทั้งสองนั้นไม่สัมพันธ์กัน เป็นต้น

แต่ถ้าค่า r จากการวิเคราะห์ เท่ากับหรือมากกว่าค่าวิกฤต แสดงว่ามีนัยสำคัญ ซึ่งถือว่าตัวแปรทั้งสองสัมพันธ์กัน (ณ ระดับนัยสำคัญนั้น) ตัวอย่างเช่น ถ้า r จากการวิเคราะห์เท่ากับ .59 ค่าวิกฤตเท่ากับ .50 ค่า r จากการวิเคราะห์มากกว่าค่าวิกฤต จึงสรุปได้ว่า ตัวแปรทั้งสองสัมพันธ์กัน

ค่าวิกฤต หาได้จากการเปิดตารางค่าวิกฤตของสัมประสิทธิ์สหสัมพันธ์โดยนำค่า df (Degree of Freedom) ซึ่งเท่ากับ $N - 2$ (เมื่อ N แทนจำนวนกลุ่มตัวอย่างทั้งหมด) ไปเปิดตาราง ซึ่งโดยทั่วไปมีระดับนัยสำคัญ 2 ระดับให้เลือก คือ ระดับนัยสำคัญ (Level of Significance : α) .05 กับระดับ .01

ระดับนัยสำคัญ .05 ซึ่งถือว่า เมื่อผู้วิจัยสรุปผลการทดสอบสมมุติฐาน ว่า ค่าเฉลี่ยหลังทดลองแตกต่างจากค่าเฉลี่ยก่อนทดลอง หรือ สูงกว่าค่าเฉลี่ยก่อนทดลอง หรือค่าเฉลี่ยหลังทดลองของกลุ่มทดลอง แตกต่างจากค่าเฉลี่ยของกลุ่มควบคุม หรือสูงกว่าค่าเฉลี่ยของกลุ่มควบคุม การสรุปเช่นนี้มีโอกาสผิดได้ 5 ใน 100 ซึ่งถือว่าน้อยมาก ผู้วิจัยจึงมีความมั่นใจในความถูกต้องร้อยละ 95 ซึ่งถือว่าสูงมากทีเดียว

ระดับนัยสำคัญ .01 ซึ่งถือว่าโอกาสสรุปผิดมี 1 ใน 100 ซึ่งน้อยกว่า 5 ใน 100 ความมั่นใจในความถูกต้องร้อยละ 99 มากกว่าระดับ .05

เมื่อสรุปว่าตัวแปรทั้งสองสัมพันธ์กัน ความสัมพันธ์ดังกล่าวอยู่ในระดับมากหรือน้อย ไม่อาจกำหนดเกณฑ์ตายตัว เช่น กำหนดว่าค่า r อยู่ระหว่าง .60 ถึง .80 หมายถึง สัมพันธ์กันมาก ทั้งนี้เนื่องจาก ระดับความสัมพันธ์ขึ้นอยู่กับกลุ่มตัวอย่างในการวิเคราะห์ครั้งนั้นว่ามีจำนวนมากหรือน้อยเป็นสำคัญ ถ้ามีจำนวนมาก เช่น 1002 คน (df เป็น 1,000) วิเคราะห์ค่า r ได้เท่ากับ .70 ก็ใช้เกณฑ์ดังกล่าวนี้ได้ (แปลว่ามีความสัมพันธ์กันมาก) แต่ถ้าจำนวนน้อย เช่น 10 คน (df เป็น 8) ค่า r .70 ไม่น่าจะสรุปว่าสัมพันธ์กันมากเพราะกรณี df เป็น 8 (โปรดดูตาราง) ค่า r จะมีนัยสำคัญที่ระดับ .05 เมื่อมีค่าเท่ากับ .632 หรือมากกว่า และเท่ากับ .765 หรือมากกว่าที่ระดับ .01 ถ้าต่ำกว่าก็ไม่มีนัยสำคัญ ซึ่งก็หมายถึงไม่สัมพันธ์กัน (ที่ระดับนั้น)

ดังนั้นถ้ามีกลุ่มตัวอย่างเพียง 10 คน เมื่อคำนวณค่า r ได้ .70 แล้ว แสดงว่าตัวแปรทั้งสองไม่สัมพันธ์กัน การกำหนดเกณฑ์ตายตัวต้องระบุจำนวนกลุ่มตัวอย่างหรือค่า df พิจารณาว่าค่าวิกฤตเป็นเท่าใด ประกอบการพิจารณาด้วย

ตาราง ค่าวิกฤตของสัมประสิทธิ์สหสัมพันธ์อย่างง่าย (r) ค่า r ที่คำนวณได้จะมีนัยสำคัญที่ระดับนั้น ถ้ามีค่ามากกว่าหรือเท่ากับค่าวิกฤตในตาราง

df (N-2)	ระดับนัยสำคัญกรณีสองหาง		df (N-2)	ระดับนัยสำคัญกรณีสองหาง	
	.05	.01		.05	.01
1	.997	1.000	24	.388	.496
2	.950	.990	25	.381	.487
3	.878	.959	26	.374	.478
4	.811	.917	27	.367	.470
5	.754	.874	28	.361	.463
6	.707	.834	29	.355	.456
7	.666	.798	30	.349	.449
8	.632	.765	35	.325	.418
9	.602	.735	40	.304	.393
10	.576	.708	45	.288	.372
11	.553	.684	50	.273	.354
12	.532	.661	60	.250	.325
13	.514	.641	70	.232	.302
14	.497	.623	80	.217	.283
15	.482	.606	90	.205	.267
16	.468	.590	100	.195	.254
17	.456	.575	125	.174	.228
18	.444	.561	150	.159	.208
19	.433	.549	200	.138	.181
20	.423	.537	300	.113	.148
21	.413	.526	400	.098	.128
22	.404	.515	500	.088	.115
23	.396	.505	1000	.062	.081

ดัดแปลงจาก R.A. Fisher and F. Yates, Statistical Tables for Biological Agricultural, and Medical Research. 6th ed.. Oliver and Boyd, Ltd, Edinburgh, 1963.

จากการพิจารณาราคาวิกฤตของสัมประสิทธิ์สหสัมพันธ์อย่างง่าย มีข้อสรุปประเด็นสำคัญดังนี้

1. เป็นตารางที่พิจารณากรณีสองหาง (Two-tailed test)
2. มีระดับนัยสำคัญที่ระดับ .05 และ .01
3. df ต่ำสุดคือ 1 แสดงถึงว่า ใช้ตารางนี้ได้เมื่อมีกลุ่มตัวอย่างน้อยที่สุด 3 คน ($df = N-2$)

อย่างไรก็ตามเมื่อมีกลุ่มตัวอย่างน้อย เช่นมีไม่ถึง 10 คน ไม่น่าศึกษาความสัมพันธ์โดยใช้วิธีการนี้

1. df สูงสุดคือ 1000 ($N = 1002$ คน) กรณีมีกลุ่มตัวอย่างมากกว่านี้ พิจารณาโดยใช้เกณฑ์ df 1000 เป็นการพิจารณาโดยประมาณ
2. ค่าวิกฤตที่ระดับ .01 จะสูงกว่าที่ระดับ .05 เสมอ เมื่อ df เดียวกัน เช่น df 50 ค่าวิกฤตที่ระดับ .01 เท่ากับ .354 สูงกว่าที่ระดับ.05 ซึ่งมีค่าวิกฤตเท่ากับ .273
3. กรณีกลุ่มตัวอย่างจำนวนมากค่า r ต่ำก็มีนัยสำคัญทางสถิติได้ แต่ถ้ามีตัวอย่างจำนวนน้อย ค่า r ต้องมีค่าสูงจึงจะมีนัยสำคัญทางสถิติ

ตัวอย่าง จำนวนกลุ่มตัวอย่าง จำนวนมาก เช่น 1002 คน ($N=1002$ $df=1000$) ค่าวิกฤตที่ระดับ .05 เท่ากับ .062 ที่ระดับ .01 เท่ากับ .081 จำนวนกลุ่มตัวอย่างน้อยเช่น 12 คน ($N=12$ $df=10$) ค่าวิกฤตที่ระดับ .05 เท่ากับ .576 ที่ระดับ .01 เท่ากับ .708

จากตัวอย่างที่ยกมาจะเห็นชัดเจนว่าเมื่อกลุ่มตัวอย่างมีจำนวนมาก ค่าวิกฤตน้อย ก็ชี้ได้ว่ามีความสัมพันธ์อย่างมีนัยสำคัญ เช่นกลุ่มตัวอย่างมี 1002 คน ค่าสหสัมพันธ์เท่ากับ .062 ก็ถือว่ามีความสำคัญทางสถิติที่ระดับ .05 ขณะที่เมื่อมีกลุ่มตัวอย่างจำนวนน้อยค่าดังกล่าวจะไม่มีนัยสำคัญทางสถิติซึ่งก็คือไม่มีสัมพันธ์กัน ตัวอย่างที่กลุ่มตัวอย่างมีจำนวนน้อยเช่นมี 12 คน ค่าสหสัมพันธ์ต้องมีค่าเท่ากับ .576 ขึ้นไปจึงจะนับว่ามีความสัมพันธ์กัน (ณ ระดับ.05) หรือที่ระดับ .01 จะพบลักษณะเช่นเดียวกัน จำนวนตัวอย่าง 1002 คน ($N=1002$ คน $df=1000$) ค่าวิกฤตเท่ากับ .081 ก็ถือว่ามีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ แต่จำนวนกลุ่มตัวอย่างมี 12 คน ค่าสหสัมพันธ์ดังกล่าว ต้องเท่ากับ .708 หรือมากกว่า จึงจะนับได้ว่ามีความสัมพันธ์กัน อย่างมีนัยสำคัญทางสถิติ

ในการพิจารณาว่าค่า r ที่คำนวณได้ ชี้ถึงว่าตัวแปรทั้งสองมีความสัมพันธ์กันอย่างไร มีนัยสำคัญหรือไม่ นอกจากจะใช้วิธีพิจารณาจากการนำค่า r ไปเปรียบเทียบกับค่าวิกฤต ดังได้กล่าวมาแล้ว ยังมีวิธีพิจารณาโดยนำค่า p ซึ่งเป็นค่า Probability ที่จะได้ค่า r ขนาดนั้นไปเปรียบเทียบกับค่า α ที่กำหนดในการวิจัยครั้งนั้น ซึ่งจะนับว่ามีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติเมื่อค่า p น้อยกว่า α วิธีนี้ยากที่จะตอบว่าสัมพันธ์กันมากเพียงใด ตอบได้พอประมาณเช่น การมีค่า p ต่ำ มากมากก็น่าจะสัมพันธ์กันมาก เป็นต้น