

บทที่ 10 สถิตินอนพารามетริกในการศึกษาความสัมพันธ์ระหว่าง ตัวแปร 2 ตัวแปร

สหสัมพันธ์

สหสัมพันธ์ (Correlation) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่ 2 ตัวขึ้นไป (หรือข้อมูล 2 ชุดขึ้นไป) ตัวอย่างการศึกษาความสัมพันธ์ เช่น การหาความสัมพันธ์ระหว่างอายุและความดันโลหิต ความสัมพันธ์ระหว่างสภาพครอบครัวกับการติดยาเสพติดในวัยรุ่น เป็นต้น สำหรับค่าสัมประสิทธิ์สหสัมพันธ์ที่เคยกล่าวถึงมาแล้ว เป็นค่าสัมประสิทธิ์สหสัมพันธ์ในสถิติพารามетริก ซึ่งได้แก่ สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน และสัมประสิทธิ์สหสัมพันธ์แบบแยกส่วน ซึ่งเป็นการหาความสัมพันธ์สำหรับตัวแปรที่มีมาตรวัดอันดับหรืออัตราส่วน และมีการแจกแจงความน่าจะเป็นแบบปกติ

ค่าสัมประสิทธิ์สหสัมพันธ์ ในสถิตินอนพารามетริก ใช้ในการหาความสัมพันธ์ระหว่างตัวแปรที่มีมาตรวัดได้ตั้งแต่นามบัญญัติขึ้นไป และไม่เจาะจงชนิดของการแจกแจงความน่าจะเป็นของข้อมูล ที่จะกล่าวถึงในบทนี้ ได้แก่

- สถิติไคสแควร์สำหรับการทดสอบความเป็นอิสระต่อกัน
(Chi-square test for independence)
- สัมประสิทธิ์สหสัมพันธ์ฟาย
- สัมประสิทธิ์สหสัมพันธ์คอนดิงเจนซี
- สัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน
- สัมประสิทธิ์สหสัมพันธ์แบบพอยท์ไบซีเรียล
- สัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล
- สัมประสิทธิ์สหสัมพันธ์เตตราคอรืค

การทดสอบไคสแควร์ สำหรับการทดสอบความเป็นอิสระต่อกัน

เป็นวิธีการที่ใช้เพื่อทดสอบสมมติฐานเกี่ยวกับความสอดคล้องของจำนวนที่อยู่ในตารางการถัว (Contingency table) กล่าวคือเป็นวิธีการที่จะพิจารณาว่าตัวแปร 2 ตัว (ที่แต่ละตัวแบ่งเป็นระดับหรือลักษณะต่างๆ นั้น) มีความเป็นอิสระต่อกัน หรือมีความสัมพันธ์กันหรือไม่ ซึ่งสถิติไคสแควร์ที่ใช้ในการทดสอบความเป็นอิสระต่อกันนี้ เป็นสถิติไคสแควร์ตัวเดียวกันกับการทดสอบไคสแควร์กรณี 2 กลุ่มอิสระต่อกัน หรือการทดสอบไคสแควร์กรณี มากกว่า 2 กลุ่มอิสระต่อกัน ดังนั้นข้อกำหนด และสถิติที่ใช้ทดสอบจึงมีความเหมือนกันทุกประการ มีข้อแตกต่างในเรื่องการตั้งสมมติฐานเท่านั้น

สมมติฐาน

H_0 : ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน (เป็นอิสระต่อกัน)

H_1 : ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน

สถิติที่ใช้ทดสอบ

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad df = (r-1)(c-1)$$

เมื่อ O_{ij} เป็นความถี่ที่ได้จากการรวบรวมข้อมูลจริงในตัวแปรที่ 1 ลักษณะที่ i และตัวแปรที่ 2

ลักษณะที่ j

E_{ij} เป็นความถี่ที่คาดว่าจะพบในตัวแปรที่ 1 ลักษณะที่ i และตัวแปรที่ 2

ลักษณะที่ j

อาณาเขตวิกฤตและการสรุปผล

จะปฏิเสธสมมติฐาน H_0 เมื่อค่า χ^2 ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า $\chi^2_{\alpha, (r-1)(c-1)}$ จากตาราง

ในที่นี้จะใช้ตัวอย่างเดียวกันกับ การทดสอบไคสแควร์กรณี 2 กลุ่มอิสระต่อกัน

ตัวอย่าง ในการวิจัยเพื่อสำรวจความวิตกกังวลในการเรียน ของนักศึกษาชายและหญิง ผลการ

สำรวจปรากฏดังตาราง

เพศ	ความวิตกกังวล	
	สูง	ต่ำ
ชาย	62	48
หญิง	73	42

จงทดสอบว่า ความวิตกกังวลในการเรียนมีความสัมพันธ์กับเพศของนักศึกษาหรือไม่

วิธีทำ

สมมติฐาน H_0 : ความวิตกกังวลในการเรียนไม่มีความสัมพันธ์กับเพศของนักศึกษา

H_1 : ความวิตกกังวลในการเรียนมีความสัมพันธ์กับเพศของนักศึกษา

สถิติที่ใช้ทดสอบ สถิติไคสแควร์

อาณาเขตวิกฤต กำหนด $\alpha = .05$ $\chi^2_{.05,1} = 3.84$ อาณาเขตวิกฤต $\chi^2 \geq 3.84$

คำนวณค่าสถิติ

เพศ	ความวิตกกังวล		รวม
	สูง	ต่ำ	
ชาย	62	48	110
	$\frac{110 \times 135}{225} = 66$	$\frac{110 \times 90}{225} = 44$	
หญิง	73	42	115
	$\frac{115 \times 135}{225} = 69$	$\frac{115 \times 90}{225} = 46$	
รวม	135	90	225

$$\begin{aligned}
 c^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(62 - 66)^2}{66} + \frac{(48 - 44)^2}{44} + \frac{(73 - 69)^2}{69} + \frac{(42 - 46)^2}{46} \\
 &= \frac{16}{66} + \frac{16}{44} + \frac{16}{69} + \frac{16}{46} = 1.186
 \end{aligned}$$

การสรุปผล

ค่า χ^2 ที่ได้จากการคำนวณ = 1.186 มีค่าน้อยกว่า ค่า χ^2 ที่เปิดจากตาราง ($\chi^2_{.05,1} = 3.84$)

ไม่ตกในอาณาเขตวิกฤต สรุปได้ว่า ความวิตกกังวลในการเรียนไม่มีความสัมพันธ์กับเพศของนักศึกษา

ผลการวิเคราะห์โดยโปรแกรมสำเร็จรูป (โปรแกรม SPSS)

GENDER * ANXIETY Crosstabulation

			ANXIETY		Total
			High	Low	
GENDER	Male	Count	62	48	110
		Expected Count	66.0	44.0	110.0
	Female	Count	73	42	115
		Expected Count	69.0	46.0	115.0
Total		Count	135	90	225
		Expected Count	135.0	90.0	225.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.186 ^b	1	.276	.281	.170
Continuity Correction ^a	.908	1	.341		
Likelihood Ratio	1.186	1	.276		
Fisher's Exact Test					
Linear-by-Linear Association	1.181	1	.277		
N of Valid Cases	225				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 44.00.

สมมติฐาน H_0 : ความวิตกกังวลในการเรียนไม่มีความสัมพันธ์กับเพศของนักศึกษา H_1 : ความวิตกกังวลในการเรียนมีความสัมพันธ์กับเพศของนักศึกษาค่า χ^2 ที่ได้จากการคำนวณ = 1.186 และมีค่า Sig. .276 > α = .05 ไม่ตกในอาณาเขต

วิกฤต สรุปว่า ความวิตกกังวลในการเรียนไม่มีความสัมพันธ์กับเพศของนักศึกษา

ตัวอย่าง ในการศึกษากลุ่มเยาวชนของชุมชนแห่งหนึ่งในปีหนึ่ง ทำการสำรวจในกลุ่มตัวอย่าง
จำนวน 200 คน ผลการสำรวจปรากฏดังตาราง

สภาพการอยู่อาศัย	การติดยาเสพติด	
	ติดยาเสพติด	ไม่ติดยาเสพติด
อยู่กับบิดา มารดา	18	165
อยู่กับญาติพี่น้อง	27	44
อยู่ตามลำพัง	105	11
รวม	150	220

จึงวิเคราะห์และสรุปผลการศึกษานี้ โดยใช้ระดับนัยสำคัญ .01

วิธีทำ

สมมติฐาน H_0 : สภาพการอยู่อาศัยของเยาวชนไม่มีความสัมพันธ์กับการติดยาเสพติด

H_1 : สภาพการอยู่อาศัยของเยาวชนมีความสัมพันธ์กับการติดยาเสพติด

สถิติที่ใช้ทดสอบ สถิติไคสแควร์

อาณาเขตวิกฤต กำหนด $\alpha = .01$ $\chi^2_{.01,2} = 9.21$ อาณาเขตวิกฤต $\chi^2 \geq 9.21$

คำนวณค่าสถิติ

สภาพการอยู่อาศัย	การติดยาเสพติด		รวม
	ติดยา	ไม่ติด	
อยู่กับบิดา มารดา	18 (74.2)	165 (108.8)	183
อยู่กับญาติพี่น้อง	27 (28.8)	44 (42.2)	71
อยู่ตามลำพัง	105 (47.0)	11 (69.0)	116
รวม	150	220	370

$$c^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(18 - 74.2)^2}{74.2} + \frac{(27 - 28.8)^2}{28.8} + \dots + \frac{(11 - 69.0)^2}{69.0}$$

$$= 191.95$$

การสรุปผล

ค่า χ^2 ที่ได้จากการคำนวณ = 191.95 มีค่ามากกว่า ค่า χ^2 ที่เปิดจากตาราง ($\chi^2_{.01,2} = 9.21$)

ตกในอาณาเขตวิกฤต สรุปได้ว่า ที่ระดับนัยสำคัญ .01 สภาพการอยู่อาศัยของเยาวชนมี
ความสัมพันธ์กับการติดยาเสพติด

ผลการวิเคราะห์โดยโปรแกรมสำเร็จรูป (โปรแกรม SPSS)

แสดงตารางการแจกแจง ระหว่าง สภาพการอยู่อาศัย (STAY) กับ ลักษณะการติดยาเสพติด (DRUG)

ข้อมูลในตารางแสดง ความถี่จากการสังเกต (Count) และความถี่คาดหวัง (Expected Count)

			DRUG		Total
			Yes	No	
STAY	1	Count	18	165	183
		Expected Count	74.2	108.8	183.0
	2	Count	27	44	71
		Expected Count	28.8	42.2	71.0
	3	Count	105	11	116
		Expected Count	47.0	69.0	116.0
Total	Count	150	220	370	
	Expected Count	150.0	220.0	370.0	

แสดงค่าสถิติไคสแควร์ (บรรทัดที่ขีดเส้นใต้) สำหรับการสรุปผลเช่นเดียวกันกับ การอ่านผลของการทดสอบภาวะรูปดี

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
<u>Pearson Chi-Square</u>	<u>191.952^a</u>	<u>2</u>	<u>.000</u>
Likelihood Ratio	214.885	2	.000
Linear-by-Linear Association	187.965	1	.000
N of Valid Cases	370		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 28.78.

สมมติฐาน H_0 : สภาพการอยู่อาศัยของเยาวชนไม่มีความสัมพันธ์กับการติดยาเสพติด H_1 : สภาพการอยู่อาศัยของเยาวชนมีความสัมพันธ์กับการติดยาเสพติดค่า χ^2 ที่ได้จากการคำนวณ = 191.95 และมีค่า Sig. .000 < α = .01 ตกในอาณาเขต

วิกฤต สรุปได้ว่า ที่ระดับนัยสำคัญ .01 สภาพการอยู่อาศัยของเยาวชนมีความสัมพันธ์กับการติดยาเสพติด

ตัวอย่างการคำนวณไคสแควร์ กรณีค่าความถี่คาดหวัง ที่ ≤ 5 เกิน 20%

ตัวอย่าง จงทดสอบว่า แหล่งที่ผู้ป่วยเข้ารับการรักษาพยาบาลเมื่อเจ็บป่วย มีความสัมพันธ์กับระดับการศึกษาของผู้ป่วยหรือไม่ จากข้อมูลที่สำรวจได้ดังนี้

ระดับการศึกษา	แหล่งที่เข้ารับการรักษา							รวม
	แพทย์แผนปัจจุบัน (คลินิก)	แพทย์แผนโบราณ	โรงพยาบาล	สถานีนามัย	ร้านขายยา	พระ	อื่น ๆ	
ไม่มีการศึกษา	41	25	19	11	23	3	7	129
ประถมศึกษา	53	9	22	6	28	1	8	127
มัธยมศึกษา	34	2	19	1	8	2	4	70
อุดมศึกษา	33	1	12	0	7	0	1	54
รวม	161	37	72	18	66	6	20	380

วิธีทำ

สมมติฐาน H_0 : แหล่งที่ผู้ป่วยเข้ารับการรักษาพยาบาลเป็นอิสระกับระดับการศึกษาของผู้ป่วย

H_1 : แหล่งที่ผู้ป่วยเข้ารับการรักษาพยาบาลมีความสัมพันธ์กับระดับการศึกษาของผู้ป่วย

ระดับนัยสำคัญ กำหนดให้ $\alpha = .01$

- การทดสอบ
- 1) สร้างตารางการนับ
 - 2) นับความถี่และบันทึกลงในแต่ละเซลล์ของตารางการนับ
 - 3) หาผลรวมในแนวแถวและผลรวมในแนวสดมภ์
 - 4) หาค่าความถี่ที่คาดหวัง (E_{ij}) ในแต่ละเซลล์ บันทึกในตารางการนับในเซลล์ O_{ij} นั้น ๆ (ในที่นี้ได้แก่ตัวเลขที่อยู่ในวงเล็บ)

ระดับการศึกษา	แหล่งที่เข้ารับการรักษา							รวม
	แพทย์แผนปัจจุบัน (คลินิก)	แพทย์แผนโบราณ	โรงพยาบาล	สถานีนานามัย	ร้านขายยา	พระ	อื่น ๆ	
ไม่มีการศึกษา	41 (54.65)	25 (12.56)	19 (24.44)	11 (6.11)	23 (22.40)	3 (2.04)	7 (6.79)	129
ประถมศึกษา	53 (53.81)	9 (12.37)	22 (24.06)	6 (6.02)	28 (22.06)	1 (2.01)	8 (6.68)	127
มัธยมศึกษา	34 (29.66)	2 (6.82)	19 (13.26)	1 (3.32)	8 (12.16)	2 (1.11)	4 (3.68)	70
อุดมศึกษา	33 (22.88)	1 (5.26)	12 (10.23)	0 (2.56)	7 (9.38)	0 (0.85)	1 (2.84)	54
รวม	161	37	72	18	66	6	20	380

จากตารางการแจกแจงดังกล่าวจะเห็นว่า มีความถี่ที่คาดหวังน้อยกว่า 5 ปรากฏถึง 8 เซลล์ ดังนั้น เพื่อให้ให้มีเซลล์ไม่ต่ำกว่า 80 เปอร์เซนต์ ที่มีความถี่ 5 หรือสูงกว่า (ตามหลักเกณฑ์ของการใช้การทดสอบไคสแควร์ ซึ่งได้กล่าวมาแล้ว) จึงควรรวมเซลล์ที่อยู่ใกล้กันหรือมีความใกล้เคียงกัน จากการรวมสมรรถที่ 3 และ 4 และสมรรถที่ 6 และ 7 เข้าด้วยกัน จะได้ผลดังตารางต่อไปนี้

ระดับการศึกษา	แหล่งที่เข้ารับการรักษา					รวม
	แพทย์แผนปัจจุบัน (คลินิก)	แพทย์แผนโบราณ	โรงพยาบาลหรือสถานีนานามัย	ร้านขายยา	พระ / อื่น ๆ	
ไม่มีการศึกษา	41 (54.56)	25 (12.56)	30 (30.55)	23 (22.40)	10 (8.83)	129
ประถมศึกษา	53 (53.81)	9 (12.37)	28 (30.08)	28 (22.06)	9 (8.69)	127
มัธยมศึกษา	34 (29.66)	2 (6.82)	20 (16.58)	8 (12.16)	6 (4.79)	70
อุดมศึกษา	33 (22.88)	1 (5.26)	12 (12.79)	7 (9.38)	1 (3.69)	54
รวม	161	37	90	66	26	380

5) หาค่า χ^2 โดยใช้สูตร

$$\begin{aligned} c^2 &= \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(41 - 54.6)^2}{54.6} + \frac{(53 - 53.81)^2}{53.81} + \dots + \frac{(1 - 3.69)^2}{3.69} \\ &= 35.786 \end{aligned}$$

การทดสอบนัยสำคัญ	นำค่า χ^2 ที่คำนวณได้ ไปเทียบกับค่าวิกฤติของ χ^2 ที่ $df = (r - 1)(c - 1) = (4 - 1)(5 - 1) = 12$ ณ $\alpha = .01$ พบว่าค่าวิกฤติของ χ^2 มีค่า 26.22
การตัดสินใจ	ค่า χ^2 ที่คำนวณได้ (35.786) มีค่ามากกว่าค่าวิกฤติของ χ^2 ที่เปิดจากตาราง (26.22) จึงปฏิเสธ H_0
แปลผล	สรุปได้ว่าแหล่งที่ผู้ป่วยเข้ารับการรักษายาบาลมีความสัมพันธ์กับระดับการศึกษาของผู้ป่วย

สัมประสิทธิ์สหสัมพันธ์ฟาย (Phi coefficient)

สัมประสิทธิ์สหสัมพันธ์ฟาย (Phi coefficient) ใช้สัญลักษณ์ ϕ เป็นวิธีที่ใช้วัดความสัมพันธ์ระหว่างตัวแปร หรือข้อมูล 2 ชุด ซึ่งเป็นข้อมูลในระดับนามบัญญัติ ผลการวัดออกมาในรูปความถี่หรือจำนวน โดยแสดงในรูปตาราง 2 X 2 เช่น การหาความสัมพันธ์ระหว่างการรัดเข็มขัดนิรภัยกับการเสียชีวิต ในกลุ่มผู้ประสบอุบัติเหตุทางรถยนต์

ข้อตกลงเบื้องต้น

ตัวแปร หรือข้อมูลทั้ง 2 ชุด มีการวัดในมาตรานามบัญญัติ และแบ่งออกเป็น 2 ลักษณะจริง (true dichotomous)

ค่าสัมประสิทธิ์สหสัมพันธ์ฟาย คำนวณจากสูตร

$$\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

เมื่อ a,b,c และ d เป็นความถี่หรือจำนวนในตาราง 2 X 2 ดังรูป

a	b
c	d

ค่า ϕ ที่ได้จะมีค่าเท่ากับ 1 ในกรณีที่ $a=d=0$ หรือ $b=c=0$

หรือ $(a+b)=(c+d)=(a+c)=(b+d)$

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0: \rho = 0$ (ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน)

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

กรณีที่ $n \geq 20$ ใช้สูตร $Z = \phi \sqrt{n}$

กรณีที่ $n < 20$ ใช้สูตร $\chi^2 = n \phi^2$; df = 1

อาณาเขตวิกฤตและการสรุปผลกรณีที่ $n \geq 20$ จะปฏิเสธ H_0 เมื่อค่า Z ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า Z ที่เปิดจากตารางกรณีที่ $n < 20$ จะปฏิเสธ H_0 เมื่อค่า χ^2 ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า χ^2 ที่เปิดจากตารางตัวอย่าง

According to data from the Department of Health and Human Services, there is an association between estrogen use and cardiovascular death rates. To corroborate these results, 1500 people were randomly selected from the population of female Americans over 50; 750 took estrogen supplements and 750 did not. The results are summarized in the following table.

		Status	
		Alive	Dead
Takes estrogen supplements?	Yes	746	4
	No	740	10

วิธีทำ

หาค่าสัมประสิทธิ์สหสัมพันธ์ จาก

$$\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}} = \frac{746(10) - 740(4)}{\sqrt{(750)(750)(1486)(14)}} \\ = \frac{4500}{108176.938} = 0.0416$$

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

 $H_0: \rho = 0$ (การใช้ estrogen กับ cardiovascular death rates ไม่มีความสัมพันธ์กัน) $H_1: \rho \neq 0$ (การใช้ estrogen กับ cardiovascular death rates มีความสัมพันธ์กัน)

กำหนดระดับนัยสำคัญ .05

กรณีที่ $n \geq 20$ ใช้ Z ค่าวิกฤต $Z = 1.96$ อาณาเขตวิกฤต $Z \leq -1.96$ หรือ $Z \geq 1.96$

สถิติทดสอบ

$$\text{ใช้สูตร} \quad Z = \phi \sqrt{n} = 0.0416 \sqrt{1500} = 1.611$$

การสรุปผล

ค่า Z ที่คำนวณ (1.611) มีค่าน้อยกว่า ค่า Z ที่เปิดจากตาราง (1.96) ไม่ตกในอาณาเขตวิกฤต ไม่ปฏิเสธ H_0 แสดงว่า การใช้ estrogen กับ cardiovascular death rates ไม่มีความสัมพันธ์กัน

ผลการวิเคราะห์โดยโปรแกรมสำเร็จรูป (โปรแกรม SPSS)

ESTROGEN * STATUS Crosstabulation

Count		STATUS		Total
		1.00	2.00	
ESTROGEN	1.00	746	4	750
	2.00	740	10	750
Total		1486	14	1500

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.042	.107
	Cramer's V	.042	.107
N of Valid Cases		1500	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

ค่า Approx. Sig. = .107 ซึ่งมากกว่า α (.05) ไม่ตกในอาณาเขตวิกฤต ไม่ปฏิเสธ H_0 แสดงว่า การใช้ estrogen กับ cardiovascular death rates ไม่มีความสัมพันธ์กัน

สัมประสิทธิ์สหสัมพันธ์คอนติเจนซี (Contingency coefficient)

สัมประสิทธิ์สหสัมพันธ์คอนติเจนซี (Contingency coefficient) ใช้สัญลักษณ์ C เป็นวิธีที่ใช้วัดความสัมพันธ์ระหว่างตัวแปร หรือข้อมูล 2 ชุด ซึ่งเป็นข้อมูลในระดับนามบัญญัติ ผลการวัดออกมาในรูปความถี่หรือจำนวน โดยแสดงในรูปตารางการแจกแจง ขนาด $r \times c$ ($r \times c$ Contingency table) ตัวอย่างเช่น การหาความสัมพันธ์ระหว่างระดับการศึกษากับการเลือกใช้สถานบริการพยาบาล การหาความสัมพันธ์ระหว่างสถานภาพสมรสกับการศึกษาต่อในระดับที่สูงขึ้นของพยาบาล

ข้อตกลงเบื้องต้น

1. ตัวแปร หรือข้อมูลทั้ง 2 ชุด มีการวัดในมาตรานามบัญญัติ หรือมีลักษณะต่อเนื่องหรือไม่ต่อเนื่องก็ได้
2. ผลการวัดอยู่ในรูปความถี่ สามารถแสดงในรูปตารางการแจกแจง ขนาด $r \times c$ (r และ $c \geq 2$)

ค่าสัมประสิทธิ์สหสัมพันธ์คอนติเจนซี คำนวณจากสูตร

$$C = \sqrt{\frac{c^2}{N + c^2}}$$

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0: \rho = 0$ (ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน)

สถิติทดสอบ เป็นการทดสอบแบบสองทาง ใช้สถิติ c^2

$$c^2 = \sum \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad df = (r-1)(c-1)$$

อาณาเขตวิกฤตและการสรุปผล

จะปฏิเสธ H_0 เมื่อค่า χ^2 ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า χ^2 ที่เปิดจากตาราง

ตัวอย่าง ในการศึกษาความสัมพันธ์ระหว่างพฤติกรรมของเด็ก กับวิธีการเลี้ยงดูในครอบครัว (3 แบบ) จากกลุ่มตัวอย่างจำนวน 60 ราย ได้ผลดังแสดง

พฤติกรรมเด็ก	วิธีการเลี้ยงดู			รวม
	แบบ 1	แบบ 2	แบบ 3	
เก็บตัว	13 (7)	4 (5.25)	4 (8.75)	21
แบบกลางๆ	5 (8)	9 (6)	10 (10)	24
แสดงตัว	2 (5)	2 (3.75)	11 (6.25)	15
รวม	20	15	25	60

วิธีทำ เมื่อทำการหาค่าความถี่คาดหวัง พบว่า $E_{ij} < 5$ มีเพียง 1 (11.11%) จึงสามารถใช้สถิติไคสแควร์ได้

หาค่าสถิติไคสแควร์จาก

$$\begin{aligned}
 \chi^2 &= \sum \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\
 &= \frac{(13 - 7)^2}{7} + \frac{(4 - 5.25)^2}{5.25} + \dots + \frac{(11 - 6.25)^2}{6.25} \\
 &= 16.87 \\
 df &= (2)(2) = 4
 \end{aligned}$$

หาค่าสัมประสิทธิ์สหสัมพันธ์คอนติเจนซีจาก

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}} = \sqrt{\frac{16.87}{60 + 16.87}} = 0.468$$

สมมติฐานของการทดสอบ

$H_0: \rho = 0$ (พฤติกรรมของเด็กกับวิธีการเลี้ยงดูในครอบครัวไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (พฤติกรรมของเด็กกับวิธีการเลี้ยงดูในครอบครัวมีความสัมพันธ์กัน)

กำหนดระดับนัยสำคัญ .05 ค่าวิกฤติ $\chi^2_{.05,4} \geq 9.49$

ค่า χ^2 ที่คำนวณได้ (16.87) มีค่ามากกว่าหรือเท่ากับค่า χ^2 ที่เปิดจากตาราง (9.49) จะปฏิเสธ H_0 แสดงว่า ที่ระดับนัยสำคัญ .05 พฤติกรรมของเด็กกับวิธีการเลี้ยงดูในครอบครัว มีความสัมพันธ์กัน

หมายเหตุ

1. ในทางทฤษฎี C จะมีค่าอยู่ระหว่าง 0 กับ 1 แต่ในทางปฏิบัติ ค่าสูงสุดของ C จะมีค่าไม่ถึง 1
2. ค่าสูงสุดของ C ขึ้นอยู่กับตารางการฉีกร หากตารางการฉีกรมีขนาดใหญ่ ค่า C จะเข้าใกล้ 1 มากขึ้น กรณีที่มีจำนวนแถวและสดมภ์เท่ากัน การประมาณค่าสูงสุดของ C หาได้จาก

$$\text{ค่าสูงสุดของ } C = \sqrt{\frac{k-1}{k}} \text{ เมื่อ } k \text{ คือจำนวนแถวและสดมภ์ที่เท่ากัน}$$

3. ในการคำนวณค่า C จะต้องคำนวณค่า χ^2 ก่อน ซึ่งก็จะมีข้อจำกัดตามสถิติ χ^2 นั้นไปด้วย
4. ค่า C ไม่สามารถนำมาใช้เปรียบเทียบกันได้โดยตรง ในขณะที่ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน แบบแยกส่วน หรือแบบสเปียร์แมน จะสามารถนำมาเปรียบเทียบกันได้โดยตรง

สัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน (Spearman rank correlation coefficient)

สัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน (Spearman rank correlation coefficient หรือ Spearman's rho) ใช้สัญลักษณ์ r_s เป็นวิธีที่ใช้วัดความสัมพันธ์ระหว่างตัวแปร หรือข้อมูล 2 ชุด โดยที่ตัวแปร หรือข้อมูล 2 ชุดนั้นจะต้องอยู่ในรูปของข้อมูลในมาตราจัดอันดับ (Ordinal scale)

ข้อตกลงเบื้องต้น

1. ตัวแปรหรือข้อมูลทั้ง 2 ชุด อยู่ในมาตราจัดอันดับ หรืออาจเป็นอันดับภาค หรือมาตราอัตราส่วน แล้วนำมาเรียงอันดับก็ได้
2. ข้อมูลในแต่ละชุดจะต้องมีความเป็นอิสระต่อกัน
สำหรับการแจกแจงของข้อมูลไม่จำเป็นต้องมีการแจกแจงแบบปกติ

ค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน คำนวณจากสูตร

$$r_s = 1 - \frac{6 \sum D^2}{N(N^2 - 1)}$$

เมื่อ r_s เป็น ค่าสัมประสิทธิ์สหสัมพันธ์แบบสเปียร์แมน

$\sum D^2$ เป็น ผลรวมของกำลังสองของผลต่างของอันดับคะแนนแต่ละคู่

N เป็น ขนาดของกลุ่มตัวอย่าง

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0: \rho = 0$ (ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน)

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

$$t = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}}, \text{ df} = n-2$$

อาณาเขตวิกฤตและการสรุปผล

จะปฏิเสธ H_0 เมื่อค่า t ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง

หรือ t ที่คำนวณได้มีค่าน้อยกว่าหรือเท่ากับค่า $-t_{\alpha, n-2}$

หรือ นำค่า r_s ที่ได้ไปเทียบกับค่าวิกฤต r_s จากตารางสำเร็จรูปโดยใช้ $\text{df} = n$

ถ้าค่า r_s มีค่ามากกว่าหรือเท่ากับ ค่าวิกฤต r_s จะปฏิเสธ H_0

ตัวอย่าง

A researcher is studying the relationship between the education level of first-born sons and the education level of their fathers in a small rural community in the Midwest . Accordingly, the researchers randomly selects 15 adults males from this community who are first-born sons and asks them to indicate on a questionnaire the highest level of education attained by themselves and by their fathers. To measure the variable, highest level of education attained, the following scale is used :

- 1 = graduated from elementary school
- 2 = graduated from middle school
- 3 = graduated from high school
- 4 = graduated from two-year college
- 5 = graduated from four-year college
- 6 = at least some graduate training

The data are as follows :

Sons	3	3	3	5	6	4	3	5	2	4	4	1	4	3	6
Fathers	1	2	1	3	3	2	3	4	3	3	2	1	2	2	4

วิธีทำ

Sons	3	3	3	5	6	4	3	5	2	4	4	1	4	3	6
Rank S.	5	5	5	12.5	14.5	9.5	5	12.5	2	9.5	9.5	1	9.5	5	14.5
Fathers	1	2	1	3	3	2	3	4	3	3	2	1	2	2	4
Rank F.	2	6	2	11	11	6	11	14.5	11	11	6	2	6	6	14.5
D	3	-1	3	1.5	3.5	3.5	-6	-2	-9	-1.5	3.5	-1	3.5	-1	0
D ²	9	1	9	2.25	12.25	12.25	36	4	81	2.25	12.25	1	12.25	1	0

หาค่าสัมประสิทธิ์สหสัมพันธ์ จาก

$$r_s = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} = 1 - \frac{6(195.5)}{15(15^2 - 1)} = 1 - \frac{1173}{3360} = 0.65$$

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0 : \rho = 0$ (Education level ของ first-born son กับ father ไม่มีความสัมพันธ์กัน)

$H_1 : \rho \neq 0$ (Education level ของ first-born son กับ father มีความสัมพันธ์กัน)

กำหนดระดับนัยสำคัญ .05

$$df = n - 2 = 15 - 2 = 13$$

ค่าวิกฤต $t_{.05,13} = 2.16$ อาณาเขตวิกฤต $t \leq -2.16$ หรือ $t \geq 2.16$

สถิติทดสอบ

$$t = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}} = \frac{.65\sqrt{15-2}}{\sqrt{1-.65^2}} = \frac{2.344}{0.760} = 3.084$$

การสรุปผล

ค่า t ที่คำนวณ (3.084) มีค่ามากกว่า ค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง (2.16) ตกในอาณาเขตวิกฤต ปฏิเสธ H_0 แสดงว่า Education level ของ first-born son กับ father มีความสัมพันธ์กันทางบวกในระดับปานกลาง

ผลการวิเคราะห์โดยโปรแกรมสำเร็จรูป (โปรแกรม SPSS)

Correlations			SONS	FATHER
Spearman's rho	SONS	Correlation Coefficient	1.000	.625*
		Sig. (2-tailed)	.	.013
		N	15	15
	FATHER	Correlation Coefficient	.625*	1.000
		Sig. (2-tailed)	.013	.
		N	15	15

*. Correlation is significant at the 0.05 level (2-tailed).

สมมติฐานของการทดสอบ

$H_0: \rho = 0$ (Education level ของ first-born son กับ father ไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (Education level ของ first-born son กับ father มีความสัมพันธ์กัน)

ค่า Sig. .013 ซึ่งน้อยกว่า α (.05) ที่ตั้งไว้ ตกในอาณาเขตวิกฤต ปฏิเสธ H_0 แสดงว่า ที่ระดับนัยสำคัญ .05 Education level ของ first-born son กับ father มีความสัมพันธ์กันทางบวกในระดับปานกลาง

สัมประสิทธิ์สหสัมพันธ์แบบพอยท์ไบเซเรียล (Point biserial correlation coefficient)

สัมประสิทธิ์สหสัมพันธ์แบบพอยท์ไบเซเรียล (Point biserial correlation coefficient) ใช้สัญลักษณ์ r_{pb} เป็นวิธีที่ใช้วัดความสัมพันธ์ระหว่างตัวแปร หรือข้อมูล 2 ชุด โดยที่ตัวแปรหนึ่งเป็นตัวแปรต่อเนื่อง อีกตัวหนึ่งมี 2 ลักษณะจริง (true dichotomous)

ข้อตกลงเบื้องต้น

1. ตัวแปรตัวหนึ่งมีการวัดอยู่ในมาตราอันตรภาค หรือมาตราอัตราส่วน และมีลักษณะการแจกแจงของประชากร ไม่จำเป็นต้องมีการแจกแจงแบบปกติ แต่ต้องมีการแจกแจงแบบโค้งเดียว และค่อนข้างสมมาตร

2. ตัวแปรอีกตัวหนึ่งมีการวัดอยู่ในมาตรานามบัญญัติและแบ่งได้เป็น 2 ลักษณะอย่างแท้จริง เช่น เพศชาย-หญิง การเป็น-ไม่เป็นโรค การตาย-รอดชีวิต คำตอบถูก-ผิด เป็นต้น

ค่าสัมประสิทธิ์สหสัมพันธ์แบบพอยท์ไบเซเรียล คำนวณจากสูตร

$$r_{pb} = \frac{\bar{X}_p - \bar{X}_q}{S_t} \sqrt{pq}$$

เมื่อ	r_{pb}	เป็น ค่าสัมประสิทธิ์สหสัมพันธ์แบบพอยท์ไบเซเรียล
	$\bar{X}_p$	เป็น ค่าเฉลี่ยของตัวแปรต่อเนื่อง ในกลุ่มลักษณะที่ 1 ของตัวแปรที่เป็น true dichotomous
	$\bar{X}_q$	เป็น ค่าเฉลี่ยของตัวแปรต่อเนื่อง ในกลุ่มลักษณะที่ 2 ของตัวแปรที่เป็น true dichotomous
	p	เป็น ค่าสัดส่วนของลักษณะที่ 1 ของตัวแปร true dichotomous
	q	เป็น ค่าสัดส่วนของลักษณะที่ 2 ของตัวแปร true dichotomous (1-p)
	S_t	ส่วนเบี่ยงเบนมาตรฐานของข้อมูลต่อเนื่องทั้งหมด

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0: \rho = 0$	(ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน)
$H_1: \rho \neq 0$	(ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน)

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

$$t = \frac{r_{pb} \sqrt{n-2}}{\sqrt{1-r_{pb}^2}}, df = n-2$$

อาณาเขตวิกฤตและการสรุปผล

จะปฏิเสธ H_0 เมื่อค่า t ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง

ตัวอย่าง จากผลการวิเคราะห์ค่าสถิติเบื้องต้นของระยะเวลาการอยู่โรงพยาบาลของคนไข้จำนวน 2469 ราย ซึ่งเป็นคนไข้เกิดแผลกดทับขณะอยู่โรงพยาบาล จำนวน 192 ราย จงหาความสัมพันธ์ระหว่างการเกิดแผลกดทับ กับระยะเวลาการอยู่โรงพยาบาล

Descriptive Statistics

bedsore		N	Minimum	Maximum	Mean	Std. Deviation
no	TODAY	2277	1	369	7.47	11.09
	Valid N (listwise)	2277				
yes	TODAY	192	1	107	16.59	18.27
	Valid N (listwise)	192				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
TODAY	2469	1	369	8.17	12.05
Valid N (listwise)	2469				

วิธีทำ

หาค่าสัมประสิทธิ์สหสัมพันธ์ จากสูตร

$$\begin{aligned}
 r_{pb} &= \frac{\bar{X}_p - \bar{X}_q}{S_t} \sqrt{pq} = \frac{16.59 - 7.47}{12.05} \sqrt{\frac{192}{2469} \times \frac{2277}{2469}} \\
 &= 0.757 \sqrt{0.0717} \\
 &= 0.203
 \end{aligned}$$

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0: \rho = 0$ (การเกิดแผลกดทับ และระยะเวลาการอยู่รพ.ของผู้ป่วยไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (การเกิดแผลกดทับ และระยะเวลาการอยู่รพ.ของผู้ป่วยมีความสัมพันธ์กัน)

กำหนดระดับนัยสำคัญ .05

$$df = n - 2 = 2469 - 2 = 2467$$

ค่าวิกฤต $t_{.05, 2467} = 1.96$ อาณาเขตวิกฤต $t \leq -1.96$ หรือ $t \geq 1.96$

สถิติทดสอบ

$$t = \frac{r_{pb} \sqrt{n-2}}{\sqrt{1-r_{pb}^2}} = \frac{0.203 \sqrt{2467}}{\sqrt{1-.203^2}} = \frac{10.083}{0.979} = 10.299$$

การสรุปผล

ค่า t ที่คำนวณ (10.299) มีค่ามากกว่า ค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง (1.96) ตกในอาณาเขตวิกฤต จึงปฏิเสธ H_0 แสดงว่า ที่ระดับนัยสำคัญ .05 การเกิดแผลกดทับ และระยะเวลาการอยู่รพ.ของผู้ป่วยมีความสัมพันธ์กัน

แสดงว่า ผู้ป่วยที่มีการเกิดแผลกดทับ มีแนวโน้มที่จะอยู่โรงพยาบาลนานกว่า ผู้ป่วยที่ไม่มีการเกิดแผลกดทับ

หมายเหตุ

ค่า r_{pb} มีค่าบวก แสดงว่าตัวอย่างในกลุ่ม p มีแนวโน้มที่จะได้คะแนนมากกว่ากลุ่ม q

ค่า r_{pb} มีค่าลบ แสดงว่าตัวอย่างในกลุ่ม p มีแนวโน้มที่จะได้คะแนนน้อยกว่ากลุ่ม q

สัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล (Biserial correlation coefficient)

สัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล (Biserial correlation coefficient) ใช้สัญลักษณ์ r_b เป็นวิธีที่ใช้วัดความสัมพันธ์ระหว่างตัวแปร หรือข้อมูล 2 ชุด โดยที่ตัวแปรหนึ่งเป็นตัวแปรต่อเนื่อง อีกตัวหนึ่งถูกแบ่งให้มี 2 ลักษณะ (force dichotomous)

ข้อตกลงเบื้องต้น

1. ตัวแปรตัวหนึ่งมีการวัดอยู่ในมาตราอันดับภาค หรือมาตราอัตราส่วน และมีลักษณะการแจกแจงของประชากร ไม่จำเป็นต้องมีการแจกแจงแบบปกติ แต่ต้องมีการแจกแจงแบบโค้งเดียว และค่อนข้างสมมาตร

2. ตัวแปรอีกตัวหนึ่งเป็นตัวแปรต่อเนื่อง แต่ถูกแบ่งออกเป็น 2 ลักษณะ โดยเกณฑ์ใดเกณฑ์หนึ่ง เช่น คะแนนความวิตกกังวล อาจแบ่งออกเป็น ความวิตกกังวลสูง-ต่ำ คะแนนภาวะสุขภาพ อาจแบ่งเป็น สุขภาพดี-ไม่ดี เป็นต้น ลักษณะการแจกแจงของประชากร จะต้องมีการแจกแจงเป็นโค้งปกติ

ค่าสัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล คำนวณจากสูตร

$$r_b = \frac{\bar{X}_p - \bar{X}_q}{S_t} \times \frac{pq}{y}$$

เมื่อ	r_b	เป็น ค่าสัมประสิทธิ์สหสัมพันธ์แบบไบซีเรียล
	$\bar{X}_p$	เป็น ค่าเฉลี่ยของตัวแปรต่อเนื่อง ในกลุ่มลักษณะที่ 1 ของตัวแปรที่เป็น force dichotomous
	$\bar{X}_q$	เป็น ค่าเฉลี่ยของตัวแปรต่อเนื่อง ในกลุ่มลักษณะที่ 2 ของตัวแปรที่เป็น force dichotomous
	S_t	ส่วนเบี่ยงเบนมาตรฐานของข้อมูลต่อเนื่องทั้งหมด
	p	เป็น ค่าสัดส่วนของลักษณะที่ 1 ของตัวแปร force dichotomous
	q	เป็น ค่าสัดส่วนของลักษณะที่ 2 ของตัวแปร force dichotomous (1-p)
	y	เป็น ค่าความสูงของโค้งปกติ ณ จุดแบ่งระหว่าง p และ q

การทดสอบนัยสำคัญสมมติฐานของการทดสอบ

$$H_0: p = 0 \quad (\text{ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน})$$

$$H_1: p \neq 0 \quad (\text{ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน})$$

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

$$t = \frac{r_{by}\sqrt{n}}{\sqrt{pq}}, \quad df = n-2$$

เมื่อ y เป็น ค่าความสูงของโค้งปกติ ที่ใช้ในสูตรหาสัมประสิทธิ์สหสัมพันธ์
 n เป็น ขนาดของกลุ่มตัวอย่าง

อาณาเขตวิกฤตและการสรุปผล

จะปฏิเสธ H_0 เมื่อค่า t ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง

ตัวอย่าง ศึกษาความสัมพันธ์ระหว่างการประเมินความเสี่ยง (อยู่ในรูปคะแนน แล้วนำมาจัดเป็น 2 กลุ่มคือ low – high risk) กับจำนวนแผลกดทับในผู้ป่วย ได้ค่าสถิติดังแสดง

Descriptive Statistics

RISKLEV		N	Minimum	Maximum	Mean	Std. Deviation
less eq 16	bedsorenumber	168	1	7	1.85	1.16
	Valid N (listwise)	168				
> 16	bedsorenumber	24	1	3	1.13	.45
	Valid N (listwise)	24				

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
bedsorenumber	192	1	7	1.76	1.12
Valid N (listwise)	192				

วิธีทำ

หาค่าสัมประสิทธิ์สหสัมพันธ์

$$\text{หาค่า } y \text{ จาก } p = \frac{168}{192} = 0.875$$

เปิดตาราง Normal Curve ที่ Area = 0.875 ได้ค่า Y ordinate = 0.2059

$$\begin{aligned}
 r_b &= \frac{\bar{X}_p - \bar{X}_q}{S_t} \times \frac{pq}{y} = \frac{1.85 - 1.13}{1.12} \times \frac{\frac{168}{192} \times \frac{24}{192}}{0.2059} \\
 &= 0.643(0.531) \\
 &= 0.341
 \end{aligned}$$

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0: \rho = 0$ (การประเมินความเสี่ยง และจำนวนแผลกดทับของผู้ป่วยไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (การประเมินความเสี่ยง และจำนวนแผลกดทับของผู้ป่วยมีความสัมพันธ์กัน)

กำหนดระดับนัยสำคัญ .05

$$df = n - 2 = 192 - 2 = 190$$

ค่าวิกฤต $t_{.05, 190} = 1.96$ อาณาเขตวิกฤต $t \leq -1.96$ หรือ $t \geq 1.96$

สถิติทดสอบ

$$t = \frac{r_b y \sqrt{n}}{\sqrt{pq}} = \frac{.341(.2059)\sqrt{192}}{\sqrt{\frac{168}{192} \times \frac{24}{192}}} = \frac{0.973}{0.331} = 2.939$$

การสรุปผล

ค่า t ที่คำนวณ (2.939) มีค่ามากกว่า ค่า $t_{\alpha, n-2}$ ที่เปิดจากตาราง (1.96) ตกในอาณาเขตวิกฤต จึงปฏิเสธ H_0 แสดงว่า ที่ระดับนัยสำคัญ .05 การประเมินความเสี่ยง และจำนวนแผลกดทับของผู้ป่วยมีความสัมพันธ์กัน ในระดับต่ำ

หมายเหตุ

- ค่า r_b ที่ได้เป็นค่าประมาณของสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน
- ในบางครั้ง ค่า r_b อาจมากกว่า 1 ได้ เนื่องจากความคลาดเคลื่อนที่เกิดจากตัวแปรที่ศึกษา มีการแจกแจงไม่ปกติ
- ค่าที่ใช้แทน ค่า r_b ได้ดีกว่าคือ r_{pb}

สัมประสิทธิ์สหสัมพันธ์เตตราคอห์ริก (Tetrachoric coefficient)

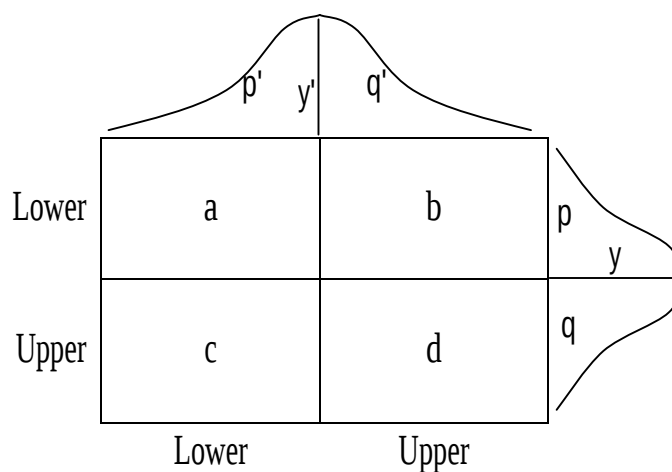
สัมประสิทธิ์สหสัมพันธ์เตตราคอห์ริก (Tetrachoric coefficient) ใช้สัญลักษณ์ r_t เป็นวิธีที่ใช้วัดความสัมพันธ์ระหว่างตัวแปร หรือข้อมูล 2 ชุด ซึ่งเป็นข้อมูลต่อเนื่อง แต่ถูกแบ่งให้มี 2 ลักษณะ (force dichotomous) ตัวอย่างเช่น การแบ่งอายุผู้ป่วยเป็นอายุต่ำกว่า 45 ปี กับ 45 ปีขึ้นไป การแบ่งคะแนนความวิตกกังวลออกเป็นกลุ่มวิตกกังวลสูง-ต่ำ เป็นต้น เมื่อข้อมูลมีการแบ่งลักษณะนี้ จะทำให้ผลการวัดออกมาในรูปความถี่หรือจำนวน โดยแสดงในรูปตาราง 2 X 2 (four fold)

ข้อตกลงเบื้องต้น

ตัวแปร หรือข้อมูลทั้ง 2 ชุด เป็นข้อมูลต่อเนื่อง มีลักษณะการแจกแจงแบบปกติ
ตัวแปร หรือข้อมูลแต่ละชุดจะถูกแบ่งออกเป็น 2 ลักษณะ

ค่าสัมประสิทธิ์สหสัมพันธ์เตตราคอห์ริก คำนวณจากสูตร

$$r_t = \frac{ad}{bc} \quad \text{หรือ} \quad r_t = \frac{ad - bc}{yy'N^2} \quad (\text{Guilford \& Fruchter, p.312})$$



a,b,c และ d เป็นความถี่หรือจำนวนในตาราง

p,q เป็นค่าสัดส่วน ของตัวแปรหรือข้อมูลชุดที่ 1 เมื่อแบ่งเป็น 2 ลักษณะ

p',q' เป็นค่าสัดส่วน ของตัวแปรหรือข้อมูลชุดที่ 2 เมื่อแบ่งเป็น 2 ลักษณะ

y เป็นความสูงของโค้ง (หรือ co-ordinate) ณ จุดแบ่ง 2 ลักษณะ ของการแจกแจงในข้อมูลชุดที่ 1

y' เป็นความสูงของโค้ง (หรือ co-ordinate) ณ จุดแบ่ง 2 ลักษณะ ของการแจกแจงในข้อมูลชุดที่ 2

N เป็นจำนวนข้อมูลทั้งหมด

การทดสอบนัยสำคัญ

สมมติฐานของการทดสอบ

$H_0: p = 0$ (ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน)

$H_1: p \neq 0$ (ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน)

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

$$Z = \frac{r_t(yy')}{\sqrt{\frac{pp'qq'}{n}}}$$

อาณาเขตวิกฤตและการสรุปผล

จะปฏิเสธ H_0 เมื่อค่า Z ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า Z ที่เปิดจากตาราง