

SHEE Research

การประยุกต์ใช้วิจัยเชิงสหสัมพันธ์ (Correlational Research) ในการศึกษาวิทยาศาสตร์สุขภาพ



ดร.เกียรติยศ กุลเดชชัยชาญ

ศูนย์ความเป็นเลิศด้านการศึกษาวิทยาศาสตร์สุขภาพ
คณะแพทยศาสตร์ศิริราชพยาบาล มหาวิทยาลัยมหิดล

วิจัยเชิงสหสัมพันธ์ เป็นหนึ่งในวิธีการค้นหาความสัมพันธ์ของตัวแปรสองตัว (หรือมากกว่า) ในกลุ่มประชากรเดียวกัน หรือหาความสัมพันธ์ในตัวแปรเดียวกันแต่มาจากประชากรหลายกลุ่ม

ข้อดีของวิจัยเชิงสหสัมพันธ์ คือ ไม่ต้องควบคุมตัวแปรเหมือนวิจัยเชิงทดลองและวิเคราะห์ตัวแปรได้หลายตัวในครั้งเดียว ซึ่งหากทำวิจัยเชิงทดลองต้องควบคุมแปรแทรกซ้อน ใช้ทรัพยากรทั้งคน เวลาและเงินทุนมากกว่า

งานวิจัยเชิงสหสัมพันธ์ จะมี 3 ขั้นตอน ซึ่งแต่ละขั้นตอน มีรายละเอียดโดยสรุป ดังนี้

1. สร้างสมมติฐาน ตัวแปรในวิจัยเชิงสหสัมพันธ์จะต้องมีทฤษฎี หลักฐานตำราและงานวิจัยในอดีต บอกว่าเกี่ยวข้องกัน เมื่อระบุตัวแปรได้แล้วจะนำมาเขียนเป็นสมมติฐานที่แสดงเงื่อนไขความสัมพันธ์ระหว่าง ตัวแปรที่ชัดเจน ขั้นตอนนี้ผู้เชี่ยวชาญมีข้อเสนอแนะว่า ก่อนจะเขียนสมมติฐานขอให้เลือกตัวแปรอิสระที่สามารถ ปรับแก้ไขได้ เพื่อที่จะได้นำไปใช้ประโยชน์ได้หลังสรุปข้อมูล ไม่ควรนำตัวแปรติดตัวมาใช้ในการวิเคราะห์ เช่น เพศ สัญชาติ ภูมิภาค แต่ให้เลือกตัวแปรที่สามารถจัดกิจกรรมส่งเสริมได้ เช่น จำนวนชั่วโมงฝึกทักษะ ปฏิบัติ ระดับความรู้เกี่ยวกับการจัดสรรเวลา ความถี่ในการออกกำลังกายต่อสัปดาห์

2. เลือกวิธีเก็บข้อมูล วิจัยเชิงสหสัมพันธ์มี 3 วิธีที่นิยมใช้ในการเก็บข้อมูล วิธีแรกคือการสำรวจ โดยใช้แบบสอบถามทั้งแบบกระดาษและแบบอิเล็กทรอนิกส์ซึ่งทำให้เก็บข้อมูลได้จำนวนมากในเวลาสั้น ข้อควรระวัง คือ ต้องทำการสุ่มและรายงานคุณภาพ (ทั้งความตรงและความเที่ยง) ของแบบสอบถาม หากเลือกใช้วิธีการสำรวจ วิธีต่อมาคือการใช้ฐานข้อมูลหรือข้อมูลทุติยภูมิที่มีอยู่ ข้อดีคือ ไม่ต้องเก็บ ข้อมูลใหม่ แต่ข้อควรระวังคือ การค้นหาฐานข้อมูลที่มีตัวแปรที่เราสนใจโดยไม่ได้วางแผนล่วงหน้าถือเป็น ไปได้ยากและในฐานข้อมูลส่วนใหญ่ก็มีข้อมูลบางส่วนที่ขาดหาย หากฐานข้อมูลที่เลือกใช้มีข้อมูลขาดหาย ไปจำนวนมาก ก็ควรพิจารณาเก็บข้อมูลใหม่แทน วิธีสุดท้ายคือการสังเกตพฤติกรรมและบันทึกข้อมูลจาก เหตุการณ์ที่เกิดขึ้น เป็นวิธีการที่ดีที่สุดในการระบุความสัมพันธ์ แต่ก็ต้องใช้เครื่องมือที่ผ่านการตรวจสอบ ความตรง และพิจารณาว่าเราเข้าถึงสถานการณ์ที่สนใจได้ไหม การสังเกตก็ควรจะต้องมีคนสังเกตมากกว่า หนึ่งคนเพื่อนำมาหาความเที่ยงข้อมูล รวมทั้งรายงานผลเพื่อความน่าเชื่อถือของผลการวิจัย



3. เก็บข้อมูล วิเคราะห์ และสรุปผล สิ่งที่สำคัญในขั้นตอนนี้คือ ก่อนการเก็บข้อมูลจะต้องมีการคำนวณขนาดตัวอย่างขั้นต่ำที่สอดคล้องกับสถิติที่ใช้ในการวิเคราะห์ เพื่อให้ผลวิจัยอ้างอิงไปสู่ประชากรได้ โดยทั่วไปงานวิจัยเชิงสหสัมพันธ์ จะใช้สถิติสองประเภท คือ การวัดความสัมพันธ์ และวิเคราะห์ การถดถอย

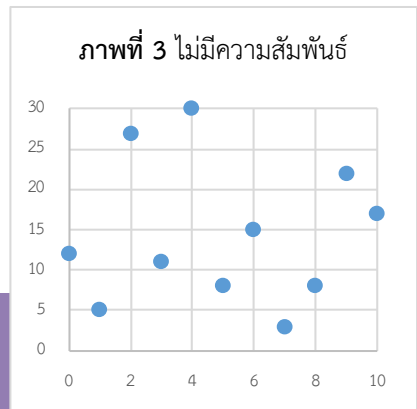
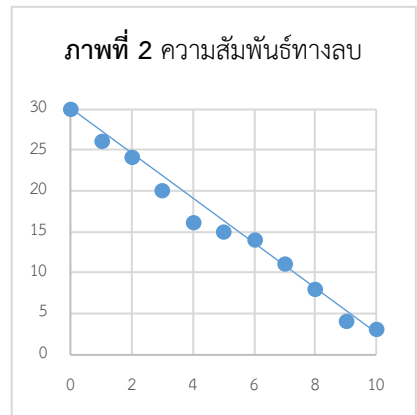
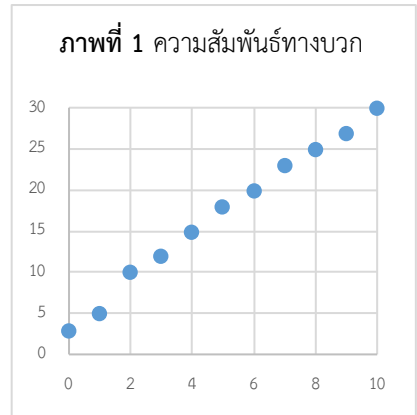
ในบทความนี้จะอธิบายถึงหลักการและวิธีแปลผลวิจัยเชิงสหสัมพันธ์เพื่อนำไปใช้เบื้องต้น ซึ่งหากผู้อ่านท่านใดยังสงสัยว่า สิ่งที่เราสนใจจะเป็นงานวิจัยเชิงสหสัมพันธ์หรือไม่ ผู้เขียนขอให้พิจารณาที่วัตถุประสงค์และคำถามวิจัย ซึ่งแบ่งเป็น 2 รูปแบบหลักคือ เพื่ออธิบายความสัมพันธ์ กับ เพื่อพยากรณ์ความสัมพันธ์ระหว่างตัวแปร ซึ่งมีรายละเอียด ดังนี้

1. เพื่ออธิบายความสัมพันธ์ระหว่างตัวแปร จะมี

คำถามวิจัย คือ “ตัวแปร ก. สัมพันธ์กับตัวแปร ข.หรือไม่ อย่างไร” ข้อมูลที่ใช้อาจนำมาจากฐานข้อมูลมีอยู่เดิมหรือเก็บข้อมูลใหม่ด้วยวิธีการสำรวจ โดยไม่ควบคุมสถานการณ์ใดๆ

วิธีการตรวจสอบความสัมพันธ์ มี 2 วิธีการ คือ **วิธีการแรก** ดูจากแผนภาพการกระจาย (Scatter diagram) หมายถึง การนำข้อมูลของตัวแปรที่สนใจมาเขียนเป็นจุดลงในกราฟเพื่อดูแนวโน้มความสัมพันธ์เมื่อลากเส้นเทียบเป็นเส้นตรง เช่น ความสัมพันธ์เป็นในทิศทางบวก คือ เมื่อตัวแปร ก. เพิ่มขึ้นจะส่งผลให้ตัวแปร ข. เพิ่มขึ้นตามภาพที่ 1 และความสัมพันธ์เป็นในทิศทางลบ คือ เมื่อ ก. เพิ่มขึ้นจะส่งผลให้ ข. ลดลงตามภาพที่ 2 หากไม่มีความสัมพันธ์ จุดจะกระจายตัวไม่สามารถลากเป็นแนวเส้นตรงได้ตามภาพที่ 3

วิธีการที่สอง ดูค่าสัมประสิทธิ์สหสัมพันธ์ (correlation coefficient : r) เป็นตัวเลขจากผลการวิเคราะห์ทางสถิติโดยนิยมใช้สูตรของเพียร์สัน ที่แสดงระดับความสัมพันธ์ระหว่างตัวแปร มีค่าตั้งแต่ -1 คือ มีความสัมพันธ์ในทิศทางลบ ถึง 1 คือ มีความสัมพันธ์ในทิศทางบวก และค่าที่เข้าใกล้ 0 คือ ไม่มีความสัมพันธ์กัน และพิจารณาค่านัยสำคัญทางสถิติ (p-value) หากต่ำกว่าระดับนัยสำคัญ 0.05 แสดงว่า ตัวแปรทั้งสองมีความสัมพันธ์กันเชิงเส้นตรง คือ มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ



2. เพื่อพยากรณ์ความสัมพันธ์ระหว่างตัวแปร จะมีคำถามวิจัย คือ “ตัวแปร ก. สัมพันธ์กับตัวแปร ข.หรือไม่ คำถามวิจัย คือ “ตัวแปรต้น ก. – ง. ทำนายตัวแปรตาม ฮ. ได้ดีเพียงใด” หรือ “สมการพยากรณ์ (ที่ประกอบด้วยตัวแปรต้นหลายตัว) ไต่บ้างที่ทำนายตัวแปรตามได้มีประสิทธิภาพ” แหล่งข้อมูลอาจใช้ฐานข้อมูลเดิม หรือสำรวจใหม่ได้เช่นกัน นิยมใช้สถิติวิเคราะห์การถดถอยหลายตัวแปร (Multiple Regression Analysis: MRA) หลักการ คือ ทดลองลากเส้นตรงผ่านจุดตัวแปรหลายตัว แล้วดูสัมพันธ์ แต่เราไม่สามารถลากเส้นผ่านทุกจุดได้ที่เก็บข้อมูลมาได้ การลากเส้นถดถอยจึงมีหลายเส้น เส้นที่ดีที่สุดจึงไม่ใช่เส้นที่ผ่านตัวแปรเยอะที่สุด แต่เป็นเส้นที่มีความคลาดเคลื่อนน้อยที่สุด ตามวิธีการกำลังสองน้อยที่สุด (the least square method) ดังนั้น ผู้อ่านต้องดูค่าพยากรณ์สำคัญทั้ง 3 กลุ่มเพื่อการนำไปใช้ คือ

- 1) **ค่าสัมประสิทธิ์การถดถอยทั้งในรูปคะแนนดิบ (b) และคะแนนมาตรฐาน (beta)** ผู้อ่านจะเปรียบเทียบได้ว่าตัวแปรต้นตัวใดส่งผลต่อตัวแปรตามมากที่สุดจากการดูค่าคะแนนมาตรฐาน (beta) รวมถึงผลทดสอบนัยสำคัญทางสถิติ จากตัวอย่างผู้วิจัยสนใจว่า ตัวแปรต้น ได้แก่ รายได้ครอบครัว นักศึกษา จำนวน ชั่วโมงอ่านบททวน และระดับความเครียดของนักศึกษา ตัวแปรใดจะสามารถทำนายคะแนนสอบ ซึ่งเป็นตัวตามได้ดีที่สุด ผลตามตารางด้านล่าง

Model	Unstandardized Coefficients		Standardized	t	Sig.	
	B	Std. Error	Beta			
1 (ค่าคงที่)	3.39	1.44		22.67	0.00	
	รายได้ครอบครัว	0.00	0.00	0.29	4.46	0.58
	เวลาอ่านบททวน	0.04	0.01	0.71	11.40	<u>0.01</u>
	ระดับความเครียด	-0.01	0.01	-0.29	-4.42	0.60
2 (ค่าคงที่)	2.95	0.72		40.99	0.00	
	เวลาอ่านบททวน	0.04	0.01	0.78	11.01	0.01

เมื่อตัวแปรตาม คือ คะแนนสอบ จากตารางดังกล่าว แผลผลได้ว่า เวลาอ่านบททวนเป็นตัวแปรต้นที่มีผลกับคะแนนสอบมากที่สุด เมื่อเทียบกับตัวแปรอื่นๆ ในสมการที่ 1 เท่ากับ 0.71 และเป็นตัวเดียวที่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติด้วย (p-value) เท่ากับ 0.01



2) ค่าสัมประสิทธิ์สหสัมพันธ์พหุคูณ ซึ่งเป็นผลจากตัวแปรตามหนึ่งตัวกับตัวแปรต้นหลายตัว และค่าสัมประสิทธิ์การพยากรณ์ (R Square: R^2) ซึ่งเป็นค่าร้อยละความแปรปรวนของตัวแปรตามที่สามารถอธิบายได้ด้วยตัวแปรต้น 3) ค่าสัมประสิทธิ์ที่เปลี่ยนแปลงจากการเพิ่มขึ้นของตัวแปรต้น (R Square change) ในสมการพยากรณ์ ประเด็นสำคัญ คือ หากการเพิ่มจำนวนตัวแปรต้นไม่ได้ทำให้ค่าดังกล่าวนี้เปลี่ยนแปลงก็เลือกใช้ตัวแปรต้นเท่าที่จำเป็น จึงทำให้ผู้วิจัยสามารถนำเสนอได้มากกว่า 1 สมการ ตามตารางด้านล่าง

model	R		Adjusted R Square	Std. Error of the Estimate	Change Statistics				
	R	Square			R Square Change	F Change	df1	df2	Sig F Change
1	0.78 ^a	0.61	0.60	0.13	0.61	121.25	1	78	0.00
2	0.80 ^b	0.64	0.64	0.12	0.64 (+0.03)	69.77	2	77	0.00
3	0.85 ^c	0.72	0.71	0.11	0.72 (+0.11)	64.19	3	76	0.00

- ตัวแปรอิสระ คือ เวลาอ่านบททวน
- ตัวแปรอิสระ คือ เวลาอ่านบททวน รายได้ครอบครัว
- ตัวแปรอิสระ คือ เวลาอ่านบททวน รายได้ครอบครัว ระดับความเครียด

จากตารางจะเห็นได้ว่า สมการที่ 3 มีค่า R และค่า R^2 สูงที่สุด คือ เท่ากับ 0.85 และ 0.72 แต่เมื่อเทียบกับค่า R Square change จากสมการที่ 1พบว่า สมการที่ 2 เท่ากับ 0.64 เพิ่มขึ้นเพียง +0.03 และสมการที่ 3 เท่ากับ 0.72 เพิ่มขึ้นเพียง +0.11 ซึ่งเพิ่มไม่มากนัก ดังนั้น ผู้เขียนจะนำเสนอตารางทั้ง 3 สมการ แต่จะสนใจอธิบายสมการพยากรณ์ที่ 1 เพราะการเพิ่มขึ้นของตัวแปรอีก 2 ตัว ไม่ได้ทำให้ค่าสัมประสิทธิ์ทั้งสามตัวเปลี่ยนแปลงมาก อีกทั้งตัวแปรต้นที่ใส่เข้าไปก็ไม่พบว่ามียุทธศาสตร์ทางสถิติด้วย จึงถือว่าการใช้ตัวแปรต้นมากเกินไป จึงเขียนสมการพยากรณ์ได้ ดังนี้ สมการที่ 1 คะแนนสอบ = 2.95 + 0.78 จำนวนชั่วโมงอ่านบททวน*

มาถึงตรงนี้น่าที่ผู้อ่านก็คงบอกว่า การวิจัยเชิงสหสัมพันธ์เพื่ออธิบายความสัมพันธ์ง่ายกว่าเพื่อพยากรณ์ความสัมพันธ์ระหว่างตัวแปรพอสมควร ซึ่งก็ถือเป็นความจริงส่วนหนึ่ง อย่างไรก็ตามการมีขั้นตอนซับซ้อนมากกว่าก็ทำให้ประยุกต์ใช้ได้มากกว่าเช่นกัน ทั้งนี้ขึ้นอยู่กับผู้อ่านว่ามีคำถามวิจัยที่สงสัยอย่างไร อีกทั้งเนื้อหาในบทความนี้เป็นกรณีศึกษาเพื่อการนำไปใช้แปลผลเบื้องต้น จึงมีรายละเอียดบางอย่างที่ยังไม่ได้ถูกกล่าวถึง เช่น ข้อตกลงเบื้องต้นทางสถิติและเทคนิคที่ใช้ตรวจสอบการวิเคราะห์ข้อมูล

หากผู้อ่านมีคำถามวิจัยที่เหมือนกับวิจัยเชิงสหสัมพันธ์ สนใจจะทำแต่มีความกังวลใจสามารถนัดหมายเพื่อพูดคุยออนไลน์กันได้จะครับที่ SHEE consult

<https://shee.si.mahidol.ac.th/index.php/consult>



การทำวิจัยเชิงสหสัมพันธ์ทางการศึกษาวิทยาศาสตร์สุขภาพ จะไม่ยากเกินฝันครับ ถ้าเราจับมือไว้แล้วไปด้วยกัน อาจารย์ศุภณ์ SHEE ยินดีต้อนรับทุกคนครับ