Correlation Dependent Rewrite

Arnond Sakworawich

ABSTRACT

For Psychological and Educational Measurement, comparing two or more dependent correlations is frequently performed, for example, to compare the predictive validity between two tests, if one test is shorter and cheaper, and also yields equal predictive validity, thus we prefer the latter, in such a task, we have to test hypothesis regarding two or more dependent correlations. The purpose of this article is to represent the newer and better three formula proposed by many researchers and statisticians. The first case is to comparing two correlations with same criterion, unfortunately, the commonly known and cited formula in many classics psychological statistics books, have been criticized that it is inexact and base on highly restrictive assumptions, the new one can avoid some fallacy with the simple calculation, the second case is to testing of heterogeneity of a set of correlated correlations, which can control type I error rate better than the traditional method, and the last case is to comparing two correlations, when sample are not independent, in most statistical textbooks represent only formula when sample are independent -using Fisher Z -Transformation, the efficiency and practicality of the test statistics are discussed.

Journal of Research Methodology, Volume 16, Number 3 (July - September 2003)

วารสารวิธีวิทยาการวิจัย ปีที่ 16 ฉบับที่ 3 (กรกฎาคม - กันยายน 2546)

การทดสอบค่าสหสัมพันธ์ที่ไม่เป็นอิสระแก่กัน

อานนท์ ศักดิ์วรวิชญ์

บทคัดย่อ

สำหรับจิตวิทยาและการวัดผลทางการศึกษาแล้ว การเปรียบเทียบค่าสหสัมพันธ์ที่ไม่เป็นอิสระแก่กันเป็นสิ่งที่ทำอยู่บ่อยครั้ง เช่นเมื่อต้องการเปรียบเทียบความตรงเชิงพยากรณ์ของแบบทดสอบสองฉบับว่าแบบทดสอบใดจะดีกว่ากัน หากแบบทดสอบใดสั้นกว่า มีค่าใช้จ่ายที่ถูกกว่าและมีความตรงเชิงพยากรณ์ที่ไม่แตกต่างกัน ก็ควรที่จะใช้แบบทดสอบนั้น ซึ่งต้องทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์ ดังนั้นบทความนี้จึงมุ่งนำเสนอสถิติทดสอบใหม่สามตัวที่มีผู้เสนอมาสำหรับการดังกล่าวโดยกรณีแรกจะเป็นการนำเสนอเปรียบเทียบค่าสหสัมพันธ์สองตัวที่มีตัวแปรเกณฑ์ตัวเดียวกัน สถิติทดสอบที่ใช้กันแพร่หลายแต่เดิมเนื่องจากพบกันในตำราสถิติทางจิตวิทยาจำนวนมากนั้น ได้มีผู้ค้นพบว่าขาดความแม่นยำและตั้งอยู่บนข้อตกลงเบื้องต้นที่เคร่งครัด ขณะที่สูตรใหม่ที่นำเสนอสามารถหลีกเลี่ยงข้อผิดพลาดบางประการโดยมีการคำนวณที่ง่าย กรณีที่สองเสนอการทดสอบความเป็นวิวิธพันธ์ของชุดสหสัมพันธ์ที่มีตัวแปรเกณฑ์ตัวเดียวกัน ซึ่งสามารถควบคุมข้อผิดพลาดแบบที่ 1 ได้ดีกว่าวิธีการแบบเดิม และกรณีสุดท้ายจะเป็นการทดสอบเปรียบเทียบค่าสหสัมพันธ์สองค่าที่กลุ่มตัวอย่างไม่เป็นอิสระแก่กัน ซึ่งปกติตำราสถิติจะเสนอสูตรที่กรณีตัวอย่างเป็นอิสระแก่กัน ซึ่งใช้การแปลงค่าซีของฟิเชอร์ ประสิทธิภาพและการนำไปใช้จะได้รับการอภิปรายต่อไป

โดยปกติเมื่อเราทดสอบค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายเพื่ออนุมานค่าสถิติ r ไปยังค่าพารามิเตอร์ ρ โดยปกติเราจะใช้สถิติทดสอบ t-test ในกรณีประชากรเดียว แต่ในกรณีที่มีสองประชากร เช่นเราต้องการทดสอบสมมุติฐานว่า ค่าสหสัมพันธ์ในสองประชากรนี้เท่ากัน หรือไม่โดยที่ค่าสหสัมพันธ์สองค่านี้เป็นอิสระแก่กัน เซอร์ อาร์ เอ ฟิเชอร์ (R.A. Fisher) เป็นผู้ที่คิดแปลงค่าสหสัมพันธ์ให้มีการแจกแจงปกติ เรียกว่าการแปลงค่าซีของฟิเชอร์ (Fisher Z-Transformation) และทำการทดสอบโดยใช้ Z-test ซึ่งสามารถพบได้ตามตำราสถิติทางจิตวิทยาหลาย ๆ เล่ม เช่น Guilford (1978), Glass & Hopkins (1984), และ Glass & Hopkins (1996) ในขณะที่หากมีการทดสอบสมมุติฐานเกี่ยวกับค่าสหสัมพันธ์มากกว่าสามประชากรที่เป็นอิสระแก่กัน จะใช้ไคกำลังสอง (χ^2) ทดสอบโดยสูตรที่เสนอโดย Hodges & Olkin (1985) ซึ่ง Glass & Hopkins (1996) กล่าวว่าแตกต่างเพียงเล็กน้อยจากสูตรเดิมที่ได้นำเสนอไว้ใน Glass & Hopkins (1984) แต่มีวิธีการคำนวณได้ง่ายกว่าและเป็นวิธีที่ดีกว่าอีกวิธีที่นำเสนอโดย Alexander, Scozzaro, & Borodkin (1990) อย่างไรก็ตามสูตรเหล่านี้จะใช้ในกรณีค่าสหสัมพันธ์ที่เป็นอิสระจากกันเท่านั้น

แต่ในกรณีที่กลุ่มตัวอย่างไม่ได้เป็นอิสระจากกันเช่นกรณีวัดซ้ำ หรือในกรณีที่กลุ่มตัวอย่างมีลักษณะร่วมกันเช่น ต้องการเปรียบเทียบความสัมพันธ์ระหว่างการเข้าชั้นเรียนกับผลการเรียนของนักเรียนสายวิทยาศาสตร์และสายศิลปะ ซึ่งต่างก็มีลักษณะร่วมคือการอยู่ในโรงเรียนเดียวกัน อาจมีบรรยากาศในการเรียนเหมือนกัน ครูผู้สอนหลาย ๆ วิชา อาจเป็นคนเดียวกัน มีวิธีสอบคัดเลือกแบบเดียวกัน ในกรณีเช่นนี้ กลุ่มตัวอย่างสองกลุ่มจะไม่เป็นอิสระแก่กัน หรือมีตัวแปรเกณฑ์ตัวหนึ่งยกตัวอย่างเช่น ผลการปฏิบัติงาน กับเชาวน์ปัญญาที่วัดจากแบบทดสอบสองฉบับเช่น APM และ WAIS และเราต้องการเลือกว่าจะเหมาะสมที่จะใช้แบบทดสอบใดมากกว่ากันกรณีเช่นนี้ เป็นกรณีที่พบได้มากในการวิจัยทางพฤติกรรมศาสตร์ แต่การใช้สถิติทดสอบค่าสหสัมพันธ์ที่ได้กล่าวไปนั้นคงไม่เหมาะแน่เพราะมีข้อสมมุติ (Assumptions) ที่ต่างไป

ในบทความนี้จะได้นำเสนอสถิติสำหรับการทดสอบสมมุติฐานเกี่ยวกับค่าสหสัมพันธ์สองค่าหรือมากกว่าที่ไม่เป็นอิสระแก่กันซึ่งพอจะแยกได้เป็น 3 กรณี คือ 1) การทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่มีตัวแปรเกณฑ์ตัวเดียวกัน 2) การทดสอบความเป็นวิวิธพันธ์ของชุดสหสัมพันธ์ที่มีตัวแปรเกณฑ์ตัวเดียวกัน และ 3) การทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่กลุ่มตัวอย่างไม่ได้เป็นอิสระแก่กันเช่น กรณีการวัดซ้ำเป็นต้น โดยจะได้ยกตัวอย่างและอธิบายตามลำดับดังนี้

1. การทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่มีตัวแปรเกณฑ์ตัวเดียวกัน

ดังที่ได้กล่าวไปแล้วในว่าการทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่มีตัวแปร

เกณฑ์ตัวเดียวกันสองตัวเป็นกรณีหนึ่งที่นักจิตวิทยาและนักวัดผลใช้กันมาก เช่นต้องการเปรียบเทียบความตรงของแบบทดสอบสองฉบับหรือการสร้างแบบทดสอบคู่ขนานที่ต้องมีความเที่ยงและความตรงใกล้เคียงกัน หรือทดสอบค่าสหสัมพันธ์สองค่าที่ขึ้นแก่กัน เช่น ต้องการเปรียบเทียบค่าสหสัมพันธ์ระหว่างบุคลิกภาพด้านการเปิดกว้างต่อประสบการณ์ของลูกกับบุคลิกภาพดังกล่าวของพ่อหรือแม่ มาเปรียบเทียบกัน เป็นต้น บางท่านอาจคิดว่าทำไมไม่พิจารณาจากค่าเบต้า (β) ซึ่งได้มาจากการวิเคราะห์ถดถอยพหุคูณ ซึ่งความจริงไม่น่าจะทำได้ดีนัก เนื่องจากในกรณีที่ตัวแปรพยากรณ์สองตัวนั้นมีความสัมพันธ์กันเองค่อนข้างสูงซึ่งทำให้เกิดปัญหาการร่วมเชิงเส้นพหุ (Multicolinearity) ซึ่งทำให้ค่า β นี้ไม่อาจแปลความได้ดีเลย ในกรณีนี้ตำราสถิติทางจิตวิทยาจำนวนมาก (Guilford,1978; Glass & Hopkins, 1984; Glass & Hopkins,1996) ได้นำเสนอสถิติทดสอบที่นำเสนอโดย Hotelling (1940) ซึ่งใช้สถิติทดสอบ t ซึ่งทดสอบ $H_0: \rho_{31} = \rho_{32}$ โดยสถิติทดสอบคือ

$$t = (r_{31} - r_{32})\sqrt{((n-3)(1+r_{12})/(2(1-r^2_{31}-r^2_{32}-r^2_{12}+2r_{31}r_{32}r_{12}))} \; ; \; \upsilon = N-3$$

โดยที่ตัวแปรเกณฑ์ต้องมีการแจกแจงปกติและความสัมพันธ์ระหว่างตัวแปรเกณฑ์และตัวแปรพยากรณ์แต่ละตัวต้องเป็นเชิงเส้น รวมถึงมีค่าความแปรปรวนเท่า ๆ กันทุก ๆ ค่าตัวแปรเกณฑ์ (Homoscedasticity)

สูตรนี้ของ Hotelling ก็ได้รับคำวิจารณ์จำนวนมาก Steiger (1980) ได้เตือนให้นักจิตวิทยาระมัดระวังการใช้สถิติทดสอบนี้โดยเฉพาะเนื่องจากสูตรนี้มีรากฐานบนข้อสมมุติที่เคร่งครัด ซึ่ง Hotelling (1940) เองก็ได้ยอมรับในบทความต้นฉบับของเขาเองว่าจะต้องยอมสูญเสียความแม่นยำจากการตั้งอยู่บนข้อสมมุติที่เคร่งครัดคือการแจกแจงปกติของตัวแปรสามตัว (Trivariate Normal Assumption)

Steiger (1980) ได้ยกตัวอย่างกรณีที่สูตรนี้จะผิดพลาดเช่น หาก $\rho_{31} = \rho_{32} = \sqrt{5}$ และ $\rho_{32} = 0$ ทั้ง ๆ ที่ความเป็นจริงแล้วสมมุติฐานว่างนี้เป็นจริงที่ระดับนัยสำคัญ ที่เราได้ตั้งไว้ก่อน แต่สถิติทดสอบของ Hotelling นี้ จะปฏิเสธสมมุติฐานว่างนี้เสมอ ข้อวิจารณ์จากท่านอื่น ๆ ถึงความไม่เหมาะสมในการใช้สถิติทดสอบนี้ได้แก่ Williams (1959), Meng, Rosenthal, & Rubin (1992)

ในอดีตได้มีผู้พยายามพัฒนาสถิติทดสอบนี้ให้ดียิ่งขึ้นได้แก่ Williams (1959), Olkin (1967), Dunn & Clark (1969), Steiger (1980), และ Meng, Rosenthal, & Rubin (1992) แต่สูตรของ Hotelling นี้ก็ยังแพร่หลายอยู่มากในหมู่นักจิตวิทยา ดังเห็นได้จากการที่ Glass &

Hopkins (1996) เองยังคงบรรจุสูตรของ Hotelling ลงในตำราของเขา โดยที่ทั้งสองท่านนี้มีความคิดเห็นว่า แม้ว่า Meng, Rosenthal, & Rubin (1992) จะได้นำเสนอสูตรที่แม่นยำกว่าของ Hotelling เล็กน้อย แต่ความแตกต่างก็ไม่มากนัก นอกจากนี้ข้อดีที่สำคัญคือสูตรของ Hotelling สามารถคำนวณได้ง่ายกว่า นอกจากนี้ Henderickson & Collins (1970) ก็ยังพบว่าสูตรที่ Williams พัฒนาใหม่ไม่ได้แตกต่างจากของ Hotelling นัก นี่คือเหตุผลสองประการที่ Glass & Hopkins (1996) ยังคงเสนอสูตรเดิมของ Hotelling

สำหรับข้อดีของสถิติทดสอบใหม่ที่นำเสนอในบทความนี้ซึ่งจะนำเสนอสูตรและตัวอย่างของ Meng, Rosenthal, & Rubin (1992) ซึ่งขยายพัฒนามาจาก Dunn & Clark (1969) มีข้อดีกว่าสูตรเดิมหลายประการ คือ 1) ไม่ต้องตั้งอยู่บนข้อสมมุติที่อาจเป็นไปได้ยากดังสูตรของ Hotelling อันได้แก่ การแจกแจงปกติของตัวแปรสามตัว (Trivariate Normal Assumption) ซึ่งอาจไม่เป็นจริงกับข้อมูลและสูตรเองก็สูญเสียความแม่นยำไป ขณะที่สูตรที่นำเสนอนี้จะมีความแม่นยำกว่า 2) การที่มีความแม่นยำมากกว่าสูตรของ Hotelling แต่ก็ ไม่ได้ยากแก่การคำนวณจนเกินกว่าจะคำนวณได้ด้วยมือดังเช่น สูตรของ Dunn & Clark (1969) หรือสูตรอื่นๆ ที่ได้นำเสนอมาโดยท่านอื่นๆ ซึ่งบางสูตรอาจต้องใช้โปรแกรมคอมพิวเตอร์มาช่วย สำหรับสูตรของ Meng, Rosenthal, & Rubin (1992) นี้ยังไม่ต้องอาศัยความรู้เรื่องเมทริกซ์ดังสูตรบางสูตรซึ่งอาจไม่เหมาะกับบางท่านที่มีพื้นความรู้ทางคณิตศาสตร์ไม่เพียงพอ 3) สูตรนี้ไม่พบข้อผิดพลาดแบบที่ 1 ดังกรณีที่ Steiger (1980) ได้ยกตัวอย่างไว้ ซึ่งจะได้ทำการแสดงตัวอย่างในภาคผนวก

อย่างไรประเด็นดังกล่าวนี้ยังมิได้มีการศึกษาวิจัยโดยการทดลองใช้การจำลองแบบมอนติ-คาร์โล เพื่อศึกษาถึงอำนาจของการทดสอบ และความแกร่ง (Robustness) ในการฝ่าฝืนข้อตกลงเบื้องต้นต่างๆ เช่น การแจกแจงที่ไม่ใช่การแจกแจงปกติ การฝ่าฝืนข้อตกลง Homoscedasticity และควรศึกษาถึงการใช้สถิติทดสอบนี้ในลักษณะที่ข้อมูลต่างๆ กัน เช่น อัตราส่วนของ N/n และ/หรือกรณีที่ตัวแปรพยากรณ์ตั้งฉากกัน หรือสัมพันธ์กันเองมากๆ ซึ่งอาจมีทั้งผลตรงและผลปฏิกิริยา หากสูตรของท่านใดที่คำนวณได้ง่ายและให้ผลดีที่ไม่ต่างกันมากนักก็น่าจะเลือกใช้สถิติทดสอบนั้นๆ ต่อไปได้ ซึ่งควรเป็นประเด็นที่นักสถิติหรือนักวัดผลน่าจะได้ลองศึกษาวิจัยกันต่อไป

สำหรับสูตรของ Meng, Rosenthal, & Rubin (1992) มีดังนี้

$$Z = (Z_{r1} - Z_{r2}) \frac{\sqrt{n-3}}{\sqrt{2(1-r_x)h}}$$

โดยที่ n = ขนาดตัวอย่าง Z_{r1} และ Z_{r2} คือ Fisher' Z Transformation ของ r_{yx1}, r_{yx2} ตามลำดับซึ่งสามารถเปิดได้จากตารางสถิติทั่วๆ ไป r_x คือค่าสหสัมพันธ์ระหว่างตัวแปรพยากรณ์

สองตัว $r_{x_1 x_2}$ นอกจากนี้

$$h = \frac{1 - f\bar{r}^2}{1 - \bar{r}^2} = 1 + \frac{\bar{r}^2}{1 - \bar{r}^2}(1 - f),$$

$$f = \frac{1 - r_x}{2(1 - \bar{r}^2)}$$

และ f ต้อง ≤ 1 และ $\bar{r}^2$ คือค่าเฉลี่ยของค่าสหสัมพันธ์กับตัวแปรเกณฑ์ทั้งสองค่า หรือ $(r_1^2 + r_2^2)$ สำหรับช่วงแห่งความเชื่อมั่นสำหรับความแตกต่างของค่า Z_rS สามารถประมาณได้ดังนี้

$$Z_{r_1} - Z_{r_2} \pm Z_{1-(\alpha/2)} \frac{\sqrt{2(1-r_x)h}}{\sqrt{N-3}}$$

ตัวอย่างกรณีที่ 1

นักจิตวิทยาอุตสาหกรรมท่านหนึ่งต้องตัดสินใจที่จะเลือกใช้แบบทดสอบซึ่งทดสอบเชาวน์ปัญญาทั้งสองฉบับโดยที่บริษัท A มีจำนวนข้อมากกว่า ใช้เวลาทดสอบมากกว่าและมีราคาแพงกว่า ในขณะที่อีกบริษัทคือบริษัท B นำเสนอแบบทดสอบที่สั้นกว่า จำนวนข้อน้อยกว่าใช้เวลาน้อยกว่าและราคาถูกกว่ากัน อย่างไรก็ตามเขาพบว่าแบบทดสอบทั้งสองนี้มีค่าความเที่ยงเท่ากันพอดีจากข้อมูลที่ผู้ขายบอก เพื่อหาข้อมูลประกอบการตัดสินใจในการเลือกซื้อ เขาได้ทดลองโดยที่บริษัททั้งสองให้สามารถทดลองฟรีได้จำนวน 15 ชุด เขาจึงสุ่มเลือกวิศวกรออกแบบระบบซึ่งเป็นพนักงานส่วนใหญ่ของบริษัทมาทดสอบจำนวน 15 คน และหาค่าความสัมพันธ์กับผลการปฏิบัติงาน (ใช้ตัวย่อ P) ได้ผลพบว่า ความสัมพันธ์ระหว่างผลการปฏิบัติงานกับคะแนนจากแบบทดสอบของบริษัท A มีค่าเท่ากับ .51 ในขณะที่ความสัมพันธ์ระหว่างผลการปฏิบัติงานกับคะแนนจากแบบทดสอบของบริษัท B มีค่าเท่ากับ .47 และพบว่าคะแนนจากแบบทดสอบ A และ B มีความสัมพันธ์กันเองมีค่าเท่ากับ .78 หากท่านเป็นนักจิตวิทยาอุตสาหกรรมท่านนี้ท่านจะตัดสินใจอย่างไร

หลักในการตัดสินใจคือ หากแบบทดสอบทั้งสองฉบับมีความตรงตามเกณฑ์สัมพันธ์ไม่ได้แตกต่างกันก็ควรเลือกใช้แบบทดสอบฉบับที่ถูกกว่าและสั้นกว่าเราจะได้แสดงการคำนวณดังนี้

ในกรณีนี้ $N = 15$, $r_{PA} = .51$, $r_{PB} = .47$, $r_{AB} = .78$ จากการเปิดตาราง Fisher's Z Transformation ได้ค่า $Z_{r1} = .563$, $Z_{r2} = .510$ ตามลำดับ เราจะได้ตั้งสมมุติฐานดังนี้

$$H_o : \rho_{PA} = \rho_{PB}$$

$$H_a : \rho_{PA} \neq \rho_{PB}$$

กำหนด α = .05

โดยที่สถิติทดสอบคือ Z-test โดยที่ Z มีค่าเท่ากับ

$$Z = (0.563 - 0.510)\frac{\sqrt{15-3}}{\sqrt{2(1-0.78)1.27079}}$$

$$Z = 0.053\,\frac{3.464}{0.5591476} = .32834$$

$$h = \frac{1 - 0.1448321 * 0.2405}{1 - 0.2405} = \frac{0.96516787}{0.7595} = 1.27079$$

$$f = \frac{1 - 0.78}{2(1 - 0.2405)} = \frac{0.22}{1.519} = 0.1448321$$

$$\bar{r}^2 = ((.51)^2 + (.47)^2)/2 = .2405$$

ค่าสถิติ Z วิกฤติจากตารางคือ −1.96 และ 1.96 แต่ค่าสถิติทดสอบที่คำนวณได้เท่ากับ 0.32 ซึ่งไม่มากหรือน้อยไปกว่าช่วงที่ได้จากการเปิดตาราง

เราจึงไม่สามารถปฏิเสธสมมุติฐานว่างได้ และไม่อาจสรุปได้ว่าค่าสหสัมพันธ์สองค่านี้แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จากข้อมูลที่มีอยู่

ดังนั้นสำหรับนักจิตวิทยาอุตสาหกรรมซึ่งต้องตัดสินใจนั้น เนื่องจากแบบทดสอบทั้งสองฉบับมีความตรงตามเกณฑ์สัมพันธ์ที่ไม่แตกต่างกันในขณะที่แบบทดสอบของบริษัท B มีราคาถูกกว่า จำนวนข้อน้อยกว่า และใช้เวลาน้อยกว่า

นอกจากการทดสอบสมมุติฐานแล้วจากตัวอย่างนี้เราสามารถคำนวณช่วงแห่งความเชื่อมั่น 95% ได้ดังนี้

$$95\%CI = (0.563 - 0.510) \pm 1.96\,\frac{\sqrt{2(1-0.78)1.27079}}{\sqrt{15-3}}$$

$$= 0.053 \pm 1.96\,\frac{0.5591476}{3.464}$$

$$= 0.053 \pm 0.31637$$

2. การทดสอบความเป็นวิวิธพันธ์ของชุดสหสัมพันธ์ที่มีตัวแปรเกณฑ์ตัวเดียวกัน (Testing of Heterogeneity of a set of Correlated Correlations)

ในกรณีที่มีตัวแปรพยากรณ์มากกว่าสองตัวนั้น การทดสอบความเป็นวิวิธพันธ์ของชุดสหสัมพันธ์ที่มีตัวแปรเกณฑ์ตัวเดียวกัน หรืออันที่จริงเราต้องการทดสอบว่า

$$H_o : \rho_{Y1} = \rho_{Y2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots = \rho_{YK}$$

ซึ่งยังมีวรรณกรรมที่เกี่ยวข้องค่อนข้างน้อยเมื่อเปรียบเทียบกับกรณีแรก (Steiger, 1980)

อย่างไรก็ตามได้มีผู้ศึกษาและคิดค้นสูตรสถิติเพื่อกรณีนี้มาหลายท่านได้แก่ Steiger (1980), Olkin & Finn (1990) อีกทางเลือกหนึ่งคือการใช้การทดสอบสมมุติฐานโดย สถิติทดสอบ t-test ที่เสนอโดย Hotelling โดยเป็นการเปรียบเทียบรายคู่ (Pairwise Comparison) ซึ่งทำได้โดยการคำนวณด้วยมือ งานวิจัยในประเทศไทยหลายชิ้นได้เลือกใช้วิธีการนี้ในการทดสอบสมมุติฐานได้แก่งานของ สุภมาส อังศุโชติ (2544) ซึ่งเปรียบเทียบความตรงเชิงทำนาย (ทำนายเกรดเฉลี่ยสะสมในระดับปริญญาตรี) ของการปรับคะแนนเฉลี่ยสะสมระดับมัธยมศึกษาตอนปลายด้วยวิธีการหลาย ๆ วิธีการ

ในบทความนี้จะได้เสนอสูตรของ Meng, Rosenthal, & Rubin (1992) ซึ่งปรับขยายมาจากสูตรกรณีที่ 1 ที่ได้นำเสนอไปแล้ว โดยขยายมาใช้สถิติทดสอบไคกำลังสอง (χ^2)

สำหรับข้อดีของวิธีนี้ที่ดีกว่าวิธีเดิม คือ 1. เมื่อเปรียบเทียบกับสูตรที่เสนอโดย Steiger (1980), Olkin & Finn (1990) ข้อที่เห็นได้ชัดเจนมากของสูตรใหม่นี้ที่ Meng, Rosenthal, & Rubin (1992) เสนอคือคำนวณง่ายกว่าทั้งสองสูตรเดิมที่ค่อนข้างยุ่งยาก โดยเฉพาะมีการคำนวณโดยเมทริกซ์ซึ่งบางท่านอาจมีความรู้ไม่เพียงพอหรือไม่อาจทำได้ แม้แต่ Steiger เองก็ยังระบุว่าสามารถใช้โปรแกรมคอมพิวเตอร์ที่เขาเขียนเองได้โดยติดต่อไปโดยตรง ซึ่งคงไม่เหมาะสมมากนักในทางปฏิบัติ และแน่นอนว่าในโปรแกรมสำเร็จรูปทางสถิติเช่น SPSS, SAS, BMDP ยังไม่มีบรรจุคำสั่งไว้ 2. การทดสอบโดยใช้การเปรียบเทียบรายคู่ ไม่ได้เป็นการทดสอบที่ตรงกับสมมุติฐานการวิจัยนักเนื่องจากนักวิจัยต้องการทดสอบโดยรวม (Omnibus Test) ซึ่งหากมีเพียงคู่ใดคู่หนึ่งแตกต่างกันไปก็ควรจะปฏิเสธสมมุติฐานว่างแล้ว ปัญหาเช่นนี้คงคล้ายกับที่เราทดสอบค่าเฉลี่ยเลขคณิตโดยใช้การทดสอบ t-test และการวิเคราะห์ความแปรปรวน 3. การเปรียบเทียบรายคู่ โดยใช้ t-test ของ Hotelling จะทำให้เกิดข้อผิดพลาดชนิดที่ 1 ค่อนข้างมาก ซึ่งอาจแก้ไขบรรเทาเบาบางลงไปโดยการกำหนดระดับนัยสำคัญน้อยกว่าระดับที่ตั้งไว้เดิม โดยหารด้วยจำนวนคู่ที่เปรียบเทียบเช่น มีค่าสหสัมพันธ์จำนวน 5 ค่าจากจำนวนตัวแปรเกณฑ์จำนวน 5 ตัว เราจะต้อง

ทดสอบทั้งหมด 10 คู่ หรือเท่ากับ (5!) / ((5-2!)2!) เราจึงควรกำหนดค่า ∝ ในการทดสอบรายคู่ด้วยสถิติทดสอบ t-test แต่ละครั้งเท่ากับ .05/10 หรือกำหนดไว้เท่ากับ .005 นั่นเอง ซึ่งวิธีการนี้ก็ค่อนข้างอนุรักษ์นิยม (Conservative) มาก เพราะโดยแท้จริงค่า actual alpha อาจจะน้อยกว่านั้น เมื่อมีข้อผิดพลาดชนิดที่หนึ่งน้อย ข้อผิดพลาดชนิดที่สองจะผันแปรเพิ่มขึ้น และอำนาจของการทดสอบจะลดลง ตามกันไป ซึ่งเป็นสิ่งที่ไม่พึงปรารถนา 4. การเปรียบเทียบรายคู่นี้น่าจะเข้าข่ายการวิเคราะห์ภายหลัง (Post hoc Analysis) ซึ่งจริง ๆ แล้วจากการทบทวนวรรณกรรมและงานวิจัยที่เกี่ยวข้อง เช่นกรณีของสุภมาส อังศุโชติ (2544) สามารถตั้ง a priori contrast ได้ แต่วิธีเดิมจะไม่เอื้อให้ทำได้ สำหรับสูตรใหม่นี้จะตั้ง a priori contrast ได้ อันจะเป็นวิธีการที่จะทดสอบสมมุติฐานการวิจัยได้ตรงกว่า และการใช้ a priori contrast นี้น่าจะมีอำนาจของการทดสอบมากกว่า a posteriori contrast เมื่อตัวแปรอื่น ๆ คงที่

สำหรับสูตรของ Meng, Rosenthal, & Rubin (1992) ที่เลือกนำเสนอในบทความนี้ เป็นดังนี้

$$\chi^2(k-1) = \frac{(N-3)\,\Sigma_i (Z_{ri} - \bar{Z}_r)^2}{(1-r_x)h}$$

โดยมีการแจกแจงแบบ χ^2 ที่องศาอิสระเท่ากับ k-1 โดยที่ k เท่ากับจำนวนค่าสหสัมพันธ์ที่ต้องการเปรียบเทียบหรือเท่ากับจำนวนตัวแปรเกณฑ์นั่นเอง ค่า h นั้นได้นำเสนอไปในกรณีที่หนึ่งไปแล้วแต่ต้องปรับโดยที่ $\bar{r}^2_N$ จะได้มาจากการหาค่าเฉลี่ยของค่าสหสัมพันธ์กำลังสองของตัวแปรพยากรณ์ทุกตัวกับตัวแปรเกณฑ์ โดยที่ $\bar{r}^2_N = \frac{i = 1^k \Sigma r^2_{Yi}}{k}$ และ r_x = ค่ามัธยฐาน (Median) ของค่าสหสัมพันธ์ระหว่างตัวแปรพยากรณ์ที่ทดสอบนี้ ต่อไปจะได้ยกตัวอย่างการคำนวณให้ดูดังนี้

ตัวอย่างกรณีที่ 2

สมมุติว่านักจิตวิทยาอุตสาหกรรมและองค์การได้ทดลองหาความตรงเชิงลู่เข้า (Convergent Validity) ของการประเมินแบบ 360 องศา โดยได้สุ่มผู้จัดการจำนวนหนึ่งเข้าไปใน ศูนย์ประเมิน (Assessment Center) เพื่อรับการประเมินจากแบบทดสอบและ In - Basket Simulation ต่าง ๆ ซึ่งทั้งหมดเป็นการประเมินศักยภาพในการบริหารโดยรวม (Overall Managerial Competencies) จำนวนทั้งสิ้น 30 คน ในขณะเดียวกันก็มีการประเมินศักยภาพในการบริหารโดยการประเมินตามการรับรู้ของ เพื่อนร่วมงานของแต่ละคน 3 คน หัวหน้างานของแต่ละคน 1 คน และผู้ใต้บังคับบัญชาของแต่ละคน 3 คน ซึ่งจะนำคะแนนที่ได้มาหาคะแนนเฉลี่ยโดยหารด้วยจำนวนผู้ประเมิน ค่าสหสัมพันธ์ได้แสดงดังตารางข้างล่างนี้

ตารางที่ 1 ค่าสหสัมพันธ์ระหว่างคะแนนเฉลี่ยของการประเมินจากแหล่งต่าง ๆ กัน (จำนวนผู้ถูกประเมิน 30 คน)

แหล่งข้อมูลของการประเมิน	Assessment Center	Peer	Boss
1.Peer Rating	.55		
2.Boss Rating	.58	.78	
3.Subordinate Rating	.41	.50	.38

นักจิตวิทยาอุตสาหกรรมผู้นี้สงสัยว่าการประเมินตามการรับรู้ของเพื่อนร่วมงาน หัวหน้างานและผู้ใต้บังคับบัญชาจะมีความตรงเชิงลู่เข้ากับ Assessment Center ที่แตกต่างกันหรือไม่

ดังนั้นเราต้องการทดสอบว่าค่าสหสัมพันธ์อย่างง่ายของการประเมินทั้ง 3 แหล่งกับ Assessment Center จะแตกต่างกันหรือไม่ซึ่งจะได้คำนวณให้ดูดังนี้

$$H_o : \rho_{AB} = \rho_{AS} = \rho_{AP}$$

$$H_a : H_o \text{ is false}$$

$$\alpha = .05$$

สถิติทดสอบคือ χ^2 โดยที่ χ^2 สามารถคำนวณได้ดังนี้

$$N = 30, r_{AP} = .55,\ Z_{r_{AP}} = 0.618,\ r_{AB} = .58, Z_{r_{AB}} = 0.662,$$

$$r_{AS} = .41, Z_{r_{AS}} = 0.436;\ \bar{Z}_r = (0.618+0.662+0.436)/3=0.572$$

$$r_x = Median = .50;\ \bar{r}^2 = ((.55)^2 + (.58)^2 + (.41)^2)/3 = .269$$

$$f = \frac{1-.50}{2(1 - .269)} = 0.341997;\ h = \frac{1-0.341997^* .269}{1 - .263} = 1.2320597$$

$$\Sigma_i(Z_{r_i} - \bar{Z}_r)^2 = (0.618-0.572)^2 + (0.662-0.572)^2 + (0.436-0.572)^2 = 0.028712$$

$$\chi_2^2 = \frac{(30 -3)^* 0.028712}{(1-.50)^* 1.2320597} = 1.25841$$

ค่าสถิติ χ^2 วิกฤติจากตารางคือ 5.99 แต่ค่าสถิติทดสอบที่คำนวณได้เท่ากับ 1.25841 ซึ่งน้อยกว่าค่าที่ได้จากการเปิดตาราง

เราจึงไม่สามารถปฏิเสธสมมุติฐานว่างได้ และไม่อาจสรุปได้ว่าค่าสหสัมพันธ์คู่ใดคู่หนึ่งนี้แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 จากข้อมูลที่มีอยู่

หากปฏิเสธ สมมุติฐานว่างเราคงอย่าง วิเคราะห์ต่อไปว่าค่าสหสัมพันธ์คู่ใดบ้างที่ต่างกัน แต่ Meng, Rosenthal, & Rubin (1992) มิได้นำเสนอ Post hoc analysis หรือการวิเคราะห์ภายหลังในบทความของพวกเขา อย่างไรก็ตามผู้เขียนมีความเห็นว่าเป็นสิ่งที่สามารถทำได้โดยใช้ Pairwise Comparison จากสูตรในกรณีที่ 1 นั่นเอง ทางเลือกหนึ่งที่น่าจะดีกว่าการควบคุมข้อผิดพลาดชนิดที่ 1 โดยการลดค่า α ลงโดยหารด้วยจำนวนคู่ที่เปรียบเทียบ กลับน่าจะใช้วิธีของ Marascuilo (1966) ซึ่งเขาพัฒนามาจากทฤษฎีของเชฟเฟ่ได้ใช้วิธีการนี้ในการวิเคราะห์ภายหลังสำหรับสถิติทดสอบต่าง ๆ หลายตัว หนึ่งในนั้นคือการทดสอบความเป็นวิวิธพันธ์ของค่าสหสัมพันธ์เมื่อกลุ่มตัวอย่างเป็นอิสระแก่กัน อย่างไรก็ตามนี่เป็นเพียงความเห็นส่วนตัวของผู้เขียนบทความนี้เท่านั้นว่าวิธีการของ Marascuilo น่าจะนำมาประยุกต์ใช้ได้ในกรณีการทดสอบความเป็นวิวิธพันธ์ของค่าสหสัมพันธ์ที่ไม่เป็นอิสระเช่นกัน วิธีการคือเมื่อคำนวณค่าสถิติทดสอบ Z จากสูตรกรณีที่หนึ่งทีละคู่ ให้นำมาเปรียบเทียบกับเกณฑ์ $\sqrt{\chi_{k-1}^2}$ แทนโดยหากค่า $|Z| \sqrt{\chi_{k-1}^2}$ โดยค่า χ^2 นี้มาจากการเปิดตารางเช่น χ^2 ที่องศาอิสระเท่ากับ 2 ได้เท่ากับ 5.99 เราจะสรุปว่าค่าสหสัมพันธ์คู่นั้น ๆ ไม่เท่ากันก็ต่อเมื่อ $|Z| \sqrt{5.99}$, $|Z| > 2.45$ นั่นเอง

นอกจากนี้ Meng, Rosenthal, & Rubin (1992) ยังได้เสนอในกรณีที่เราสามารถทำ planned comparison โดยสร้าง contrast มาล่วงหน้าหากมีทฤษฎีมารองรับมากเพียงพอที่เราจะบอกได้นั่นเองซึ่งเป็นการทดสอบสมมุติฐานที่ตรงประเด็นมากกว่า ทั้งการใช้ a priori contrast นี้น่าจะมีอำนาจของการทดสอบมากกว่า a posteriori contrast เมื่อตัวแปรอื่น ๆ คงที่ โดยต้องวางแผน contrast ไว้ก่อนการเก็บข้อมูล

Meng, Rosenthal, & Rubin (1992) ได้เสนอให้ใช้ Z test ในการทดสอบ a priori contrast ดังนี้

$$Z - \Sigma\lambda_i Z_{r_i} \frac{\sqrt{N-3}}{\sqrt{(\Sigma\lambda_i^2)(1-r_x)_h}}$$

โดยที่ λ_i คือ contrast weight ของแต่ละค่า Z_{r_i} ตาม contrast ที่ได้วางแผนไว้แต่ต้น ซึ่ง $\Sigma\lambda_i = 0$ ในกรณีที่ k มีจำนวนมาก (มีตัวแปรพยากรณ์มาก) และมีค่า λ_i บางตัวมีค่าเป็นศูนย์

เนื่องจากไม่ได้ต้องการเปรียบเทียบเราสามารถคำนวณค่า h และ r_x จากเฉพาะค่าที่เกี่ยวข้องคือ มี contrast weights ไม่เป็นศูนย์ เช่นกรณีมี contrast weights เป็น 1,1,0,0,-2 ซึ่งมีตัวแปรพยากรณ์ที่ต้องการเปรียบเทียบเพียง 3 ตัวจากห้าตัวก็สามารถหาค่ามัธยฐานจากค่าสหสัมพันธ์ระหว่างตัวแปรพยากรณ์สามตัวนี้ได้เลย เช่นเดียวกับค่า h ที่ใช้แทนในสูตรนี้ เพื่อให้ได้ความแม่นยำมากขึ้นกว่าการใช้ค่ารวมที่มาจากพารามิเตอร์ k ตัวนั่นเอง อนึ่งเราสามารถคำนวณค่าสถิติทดสอบ contrast ดังกล่าวได้โดยง่ายหากได้มีการทดสอบความเป็นวิวิธพันธ์ของชุดสหสัมพันธ์นี้มาก่อนโดยหาค่าสหสัมพันธ์อย่างง่ายระหว่าง contrast weights กับ Z_{r_i} ที่คู่กัน เช่นในกรณีนี้ก็จะมีสามคู่และนำค่าสหสัมพันธ์ดังกล่าวมาคูณกับกรณฑ์ที่สองของค่าไคกำลังสองที่ได้จากการคำนวณไว้ข้างต้น ดังสูตรข้างล่างนี้

$$Z = r_{\lambda Z_r} \sqrt{\chi^2_{(k-1)}}$$

นอกจากนี้ยังสามารถประมาณค่าแบบช่วงจาก contrast ที่วางแผนไว้ได้โดยสูตรข้างล่างนี้

$$\Sigma \lambda_i Z_{r_i} \pm Z_{1-(\alpha/2)} \frac{\sqrt{(\Sigma\lambda_i^2)(1-r_x)h}}{\sqrt{N-3}}$$

ตัวอย่างกรณีที่ 2.1

จากตัวอย่างกรณีที่ 2 สมมุติว่ามีทฤษฎีว่าผู้ใต้บังคับบัญชาหรือเพื่อนร่วมงานประเมินจะมีความตรงน้อยกว่าผู้บังคับบัญชาประเมินอันเนื่องมาจากเพื่อนร่วมงานอาจต้องแข่งกันทำผลงาน ขณะที่ลูกน้องบางคนอาจถูกนายต่อว่าทำให้ การประเมินจากสองแหล่งนี้มีอคติมากกว่าผู้บังคับบัญชาประเมิน เราอาจจะตั้ง contrast weights ดังนี้ 1 สำหรับ r_{AP},1 สำหรับ r_{AS}, และ -2 สำหรับ r_{AB} หรือสำหรับค่าสหสัมพันธ์ของเกณฑ์กับการประเมินของเพื่อนร่วมงาน ลูกน้องและนายตามลำดับและสามารถคำนวณค่า Z ได้ดังนี้

$$H_o : \varphi = O$$

$$\varphi = \rho_{AS} + \rho_{AP} - 2\rho_{AB}$$

กำหนดให้ α=.05

สถิติทดสอบคือ Z test โดยที่ Z-test สามารถคำนวณได้ดังข้างล่างนี้

$$\Sigma \lambda_i Z_{r_i} = 1*0.618 + 1*0.436 - 2*0.662 = -0.27$$

$$Z = -0.27 \frac{\sqrt{30-3}}{\sqrt{(6)(1-.50)1.2320597}}$$

$Z = -0.7297$

หรืออาจคำนวณจากอีกวิธีก็ได้เช่นกัน โดยคำนวณค่าสหสัมพันธ์ระหว่างค่า contrast weights กับ Z_{r_i} ที่คู่กัน -.651 คูณกับค่าไคกำลังสองที่คำนวณได้ในตัวอย่างที่สอง

$$Z = -.651 * \sqrt{1.25841} = -0.7302$$

ค่าที่ต่างกันเพียงเล็ก ๆน้อย ๆ นี้เกิดจากทศนิยมที่ประมาณเท่านั้นเอง

ค่า Z ที่คำนวณได้เท่ากับ -0.73 ซึ่งตกอยู่ในช่วง (-1.96,1.96) จึงไม่ปฏิเสธสมมุติฐานว่าง

อาจสรุปได้ว่าจากข้อมูลที่มีอยู่นี้การประเมินจากหัวหน้างานต่างมีความตรงไม่แตกต่างจากการประเมินโดยลูกน้องและเพื่อนร่วมงาน

นอกจากนี้เรายังสามารถประมาณค่าแห่งความเชื่อมั่น 95 % ได้ดังนี้

$$-0.27 \pm 1.96 * \frac{\sqrt{(6)(1-.50)1.2320597}}{\sqrt{30-3}}$$

$$-0.27 \pm 0.7253 = (-0.9953, 0.4553)$$

3. การทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่กลุ่มตัวอย่างไม่ได้เป็นอิสระแก่กัน

การทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่กลุ่มตัวอย่างไม่ได้เป็นอิสระแก่กันเช่น กรณีการวัดซ้ำ (Repeated Measured) หรือข้อมูลมีลักษณะทับเกี่ยวกัน (Nested Data) หรือข้อมูลมีการจับคู่ (Matched- Pair Design) เป็นต้น ในที่นี้จะขอยกตัวอย่างเช่น กรณีวัดซ้ำ เช่น สมมุติว่านักสรีระวิทยาหาความสัมพันธ์ของน้ำหนักกับส่วนสูงเปรียบเทียบ ช่วงทารกแรกเกิดกับวัยผู้ใหญ่วัย 30 ปี โดยคาดว่าเด็กทารกจะมีค่าสหสัมพันธ์มากกว่าผู้ใหญ่ที่จะได้รับผลการเลี้ยงดูหรืออุปนิสัยในการรับประทานเข้ามาเกี่ยวข้อง ดังนั้นจึงคาดว่าค่าสหสัมพันธ์ระหว่างน้ำหนักกับส่วนสูงของเด็กจะมีค่ามากกว่าของผู้ใหญ่ การออกแบบการวิจัยเป็นการวิจัยแบบระยะยาว (Longitudinal Study) โดยตามเก็บข้อมูลแต่แรกเกิดจนอายุ 30 ปี หรือกรณีที่สองข้อมูลมีลักษณะทับเกี่ยวกัน (Nested Data) เช่น เก็บข้อมูลมาจากครอบครัวเดียวกัน เช่นความสัมพันธ์ระหว่าง BMI: Body Mass Index ซึ่งคำนวณได้จากส่วนสูงหารด้วยน้ำหนักตัวยกกำลังสอง กับความดันตัวบน SBP: Systolic Blood Pressure โดยต้องการเปรียบเทียบค่าสหสัมพันธ์ระหว่างตัวแปรทางสรีระวิทยาทั้งสอง โดยคาดว่าค่าสหสัมพันธ์ของผู้เป็นแม่น่าจะมากกว่าลูกเนื่องจากแม่มีอายุมากกว่า เป็นต้น เนื่องจากเป็นครอบครัวเดียวกันจึงไม่เป็นอิสระแก่กัน หรือในกรณีการจับคู่

(Matched- Pair Design) เช่น นักจิตวิทยาการศึกษาท่านหนึ่งสนใจทำการทดลองเปรียบเทียบผลของวิธีการสอนแบบสืบเสาะ (Inquiry Method) กับการสอนแบบบรรยาย ที่มีผลต่อผลสัมฤทธิ์ในวิชาวิทยาศาสตร์ เพื่อป้องกันผลของตัวแปรแทรกซ้อนเช่นพื้นเชาวน์ปัญญาที่ต่างกันจึงได้จับคู่ระหว่างกลุ่มทดลองกับกลุ่มควบคุมไว้ นอกจากจะสนใจที่จะเปรียบเทียบค่าเฉลี่ยของตัวแปรตามแล้ว นักจิตวิทยาการศึกษาท่านนี้ยังคาดว่า ในกลุ่มที่ใช้วิธีการสอนแบบสืบเสาะนั้นความกระหายใคร่รู้ (Curiosity) ก่อนเริ่มเรียนจะมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์มากกว่าการสอนแบบบรรยาย ทั้งนี้เนื่องจากนักเรียนในกลุ่มที่เรียนแบบสืบเสาะจะมีโอกาสได้ใช้ศักยภาพในการแสวงหาความรู้มากกว่า คนที่กระหายใคร่รู้มากก็จะได้เรียนรู้มากกว่า ในขณะที่การสอนแบบบรรยายเป็นการสอนโดยที่คนที่กระหายใคร่รู้ต่างกันก็ไม่มีผลต่างกันมากนักเนื่องจากไม่ต้องค้นคว้ามากด้วยตนเอง การเปรียบเทียบค่าสหสัมพันธ์ในกรณีทั้งหมดนี้จะใช้ Z-test ซึ่งแปลงจาก Fisher Z-Transformation ไม่ได้ เนื่องจากกลุ่มตัวอย่างทั้งสองไม่เป็นอิสระขาดจากกัน

ในกรณีเช่นนี้ได้มีผู้เสนอสถิติทดสอบสำหรับกรณีดังกล่าวดังนี้ (Olkin, 1967)

$$Z = \frac{\sqrt{N}(r_{12} - r_{34})}{\hat{\sigma}r_{12} - r_{34}}$$

$$\hat{\sigma}^2_{r12} - r_{34} = (1-r^2_{12})^2 + (1-r^2_{34})^2 + r_{12}r_{34}(r^2_{13} + r^2_{14} + r^2_{23} + r^2_{24}) + 2(r_{13}r_{24} + r_{14}r_{23})$$

$$-2(r_{12}r_{13}r_{14} + r_{12}r_{23}r_{24} + r_{13}r_{23}r_{24} + r_{14}r_{24}r_{34})$$

ต่อไปจะได้แสดงตัวอย่างให้ดูดังนี้

ตัวอย่างกรณีที่ 3

สมมุติว่านักจิตวิทยาการศึกษาท่านหนึ่งสนใจทำการทดลองเปรียบเทียบผลของวิธีการสอนแบบสืบเสาะ (Inquiry Method) (กลุ่มทดลอง) กับการสอนแบบบรรยาย (กลุ่มควบคุม) ที่มีผลต่อผลสัมฤทธิ์ในวิชาวิทยาศาสตร์ เพื่อป้องกันผลของตัวแปรแทรกซ้อนเช่น พื้นเชาวน์ปัญญาที่ต่างกันจึงได้จับคู่ระหว่างกลุ่มทดลองกับกลุ่มควบคุมไว้ นอกจากจะสนใจที่จะเปรียบเทียบค่าเฉลี่ยของตัวแปรตามแล้ว นักจิตวิทยาการศึกษาท่านนี้ยังคาดว่า ในกลุ่มที่ใช้วิธีการสอนแบบสืบเสาะนั้นความกระหายใคร่รู้ (Curiosity) ก่อนเริ่มเรียนจะมีความสัมพันธ์กับผลสัมฤทธิ์ทางการเรียนวิชาวิทยาศาสตร์มากกว่าการสอนแบบบรรยาย ทั้งนี้เนื่องจากนักเรียนในกลุ่มที่เรียนแบบสืบเสาะจะมีโอกาสได้ใช้ศักยภาพในการแสวงหาความรู้มากกว่า คนที่กระหายใคร่รู้มากก็จะได้เรียนรู้มากกว่า ในขณะที่การสอนแบบบรรยายเป็นการสอนโดยที่คนที่กระหายใคร่รู้ต่างกันก็ไม่มีผลต่างกันมากนักเนื่องจากไม่ต้องค้นคว้ามากด้วยตนเอง ค่าสหสัมพันธ์ได้แสดงไว้ในตารางข้างล่างนี้

ตารางที่ 2 แสดงค่าสหสัมพันธ์ระหว่างตัวแปรในการทดลองหนึ่งของนักจิตวิทยาการศึกษา (จำนวนพลวิจัยของกลุ่มทดลองเท่ากับกลุ่มควบคุมคือกลุ่มละ 25 คน)

ตัวแปร	1C_E	2P_E	3C_C
1 ความกระหายใคร่รู้: C_E (กลุ่มทดลอง)			
2 ผลสัมฤทธิ์วิชาวิทยาศาสตร์: P_E (กลุ่มทดลอง)	.78		
3 ความกระหายใคร่รู้: C_C (กลุ่มควบคุม)	.65	.45	
4 ผลสัมฤทธิ์วิชาวิทยาศาสตร์: P_C (กลุ่มควบคุม)	.23	.23	.42

ต่อไปจะได้แสดงการคำนวณดังนี้

$$H_o : \rho_{C_E P_E} \leq \rho_{C_T P_T}$$

$$H_a : \rho_{C_E P_E} > \rho_{C_T P_T}$$

$$\propto \ = .05$$

สถิติทดสอบคือ z-test โดยคำนวณได้ดังนี้

$$Z = \frac{\sqrt{25}\ (.78-.42)}{0.9445959982} = \frac{0.36*5}{0.9445959082} = 1.90557$$

$$\hat{\sigma}^2 r_{C_E P_E} - r_{C_C P_C} = (1-(.78)^2)^2 + (1-(.42)^2)^2 + (.78)(.42)(.65^2 + .23^2 + .45^2 + .23^2)$$

$$+ 299.65)(.23)+(.23)(.45)-2((.78)(.65)(.23)+(.78)(.45)(.23)+(.65)(.45)(.42)+ (.23)(.23)(.42)$$

$$\hat{\sigma}^2 r_{C_E P_E} - r_{C_C P_C} = 0.153 + 0.678 + 0.239 + 0.506 - 0.684 = 0.892$$

$$\hat{\sigma}^2 r_{C_E P_E} - r_{C_C P_C} = \sqrt{0.892} = 0.944$$

เนื่องจากค่าสถิติ z ที่คำนวณได้มีค่าเท่ากับ 1.905 ซึ่งมากกว่า 1.645 เราจึงปฏิเสธสมมุติฐานว่าง

แสดงว่าจากข้อมูลที่มีอยู่ค่าสหสัมพันธ์ระหว่างความกระหายใคร่รู้กับผลสัมฤทธิ์ในการเรียนวิชาวิทยาศาสตร์ของกลุ่มที่สอนโดยการเรียนรู้แบบสืบเสาะมีค่ามากกว่าแบบบรรยาย

จากที่ได้นำเสนอมาทั้งหมดสามกรณีนี้น่าจะเป็นประโยชน์บ้างตามสมควร อนึ่งการทดสอบความเป็นวิวิธพันธ์ของค่าสหสัมพันธ์หลายค่าค่าที่กลุ่มตัวอย่างไม่ได้เป็นอิสระแก่กันเช่น กรณีการวัดซ้ำ (repeated measured) หรือข้อมูลมีลักษณะทับเกี่ยวกัน (nested data) หรือข้อมูลมีการจับคู่ (matched- pair design) เช่นเดียวกับกรณีที่ 3 แต่มีค่าสหสัมพันธ์หลาย ๆ ค่าที่ต้องการทดสอบพร้อม ๆ กัน ไม่ได้นำเสนอไว้ในบทความนี้ เนื่องจากสถิติทดสอบนี้ ต้องใช้เมทริกซ์มาช่วยมากพอควร โดยหากมีค่าสหสัมพันธ์มาก ๆ หลาย ๆ ค่าที่ต้องการทดสอบ การคำนวณด้วยมือคงเป็นการพ้นวิสัย ซึ่งสูตรในกรณีนี้ ต้องมีการหาเมทริกซ์ผกผัน (inverse matrix) และหากเมทริกซ์ผกผันดังกล่าวมีขนาดใหญ่ ๆ หรือหากค่า determinant ของเมทริกซ์ดังกล่าวมีค่าเท่ากับศูนย์ ซึ่งเป็นกรณีที่เรียกว่าเมทริกซ์เอกฐาน (singular matrix) ก็ยังเป็นปัญหาสำคัญในการคำนวณหาเมทริกซ์ผกผัน (inverse matrix) อย่างไรก็ตามแม้ไม่มีโปรแกรมสำเร็จรูปทางสถิติใดที่มีการวิเคราะห์กรณีนี้ได้ โปรแกรมคำนวณทางคณิตศาสตร์บางโปรแกรมเช่น MATLAB ก็น่าจะเป็นประโยชน์อยู่มากและไม่ยากไปกว่าการโปรแกรมมิ่งเอง สำหรับผู้สนใจกรณีการทดสอบความเป็นวิวิธพันธ์ของค่าสหสัมพันธ์หลายค่าค่าที่กลุ่มตัวอย่างไม่ได้เป็นอิสระแก่กันนี้ สามารถศึกษาได้จาก Olkin & Finn, 1990 ในส่วนของ Model B ต่อไป

สรุปบทความนี้มุ่งเสนอการทดสอบค่าสหสัมพันธ์ที่ไม่เป็นอิสระแก่กันในสามกรณีคือ 1.การทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่มีตัวแปรเกณฑ์ตัวเดียวกัน 2. การทดสอบความเป็นวิวิธพันธ์ของชุดสหสัมพันธ์ที่มีตัวแปรเกณฑ์ตัวเดียวกัน และ 3. การทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่กลุ่มตัวอย่างไม่ได้เป็นอิสระแก่กันเช่น กรณีการวัดซ้ำเป็นต้น ทั้งนี้ผู้อ่านควรทำความเข้าใจเรื่องข้อสมมุติให้ชัดเจนก่อนการนำไปใช้ในการทดสอบสมมุติฐาน อนึ่งค่าสถิติทดสอบต่าง ๆ เหล่านี้ยังไม่มีบรรจุในโปรแกรมสำเร็จรูปทางสถิติทั้งสิ้น แต่ผู้เขียนเลือกนำเสนอเฉพาะสูตรที่อยู่ในวิสัยที่สามารถจะคำนวณด้วยมือได้ และหวังว่าบทความนี้จะเป็นส่วนหนึ่งที่ช่วยให้เกิดการใช้สถิติในการวิจัยได้อย่างดียิ่งขึ้น ๆ ไป และหวังจะเห็นมีการคิดค้นสถิติตัวใหม่ ๆ เพื่อให้ผลการวิเคราะห์ที่ดีกว่าเดิมหรือเหมาะสมกับการทดสอบสมมุติฐานการวิจัยมากขึ้น

เอกสารอ้างอิง

สุภมาส อังศุโชติ. (2544). การปรับคะแนนเฉลี่ยสะสมระดับมัธยมศึกษาตอนปลาย: การเปรียบเทียบความตรงเชิงทำนาย. **วารสารวิธีวิทยาการวิจัย** 14, 3: 270-293.

Alexander, R.A., Scozzaro, M.J., & Borodkin, L.J. (1990). Statistical and empirical examination of the chi-square test for homogeneity of correlations in meta-analysis. **Psychological Bulletin** 106: 329-331.

Dunn, O.J., & Clark, V.A. (1969). Correlation coefficient measured on the same individuals. **Journal of the American Statistical Association** 64: 366-377. quoted in Meng, Xiao-Li., Rosenthal, Robert., & Rubin, Donald. B. (1992). Comparing Correlated Correlation Coefficients. **Psychological Bulletin** 111 1: 172-175.

Glass, Gene V., & Hopkins, Kenneth D. (1984). **Statistical methods in education and psychology**. MA: Allyn & Bacon.

Glass, Gene V., & Hopkins, Kenneth D. (1996). **Statistical methods in education and psychology**. MA: Allyn & Bacon.

Guilford, J.P., & Fruchter, B. (1978). **Fundamental statistics in psychology and education**. New York: McGraw-Hill.

Hedges, L.V., & Olkin, I. (1985). **Statistical methods for meta-analysis**. New York: Academic Press.

Henderickson, G.F., & Collins, J.R. (1970). Note correcting the results in Olkin's new formula for the significance of r_{13} vs. r_{23} compared with Hotelling's methods. **American Educational Research Journal** 7: 639-664. quoted in Glass, Gene V., & Hopkins, Kenneth D. (1996). **Statistical methods in education and Psychology**. MA: Allyn & Bacon.

Hotelling, H. (1940). The selection of variates for use in prediction with some comments on the general problem of nuisances parameters. **Annals of Mathematical Statistics** 11: 271-283. in Meng, Xiao-Li., Rosenthal, Robert., & Rubin, Donald. B. (1992). Comparing Correlated Correlation Coefficients. **Psychological Bulletin,** 111, 1: 172-175.

Marascuilo, L.A. (1966). Large-sample multiple comparisons. **Psychological Bullettin**, 15, 5: 280-290.

Meng, Xiao-Li., Rosenthal, Robert., & Rubin, Donald. B. (1992). Comparing correlated correlation coefficients. **Psychological Bulletin** 111 1: 172-175.

Olkin, I. (1967). Correlation revisited, in **Improving experimental design and statistical analysis**, Stanley, J.C. Chicago: Rand McNally. quoted in Glass, Gene V., & Hopkins, Kenneth D. (1996). **Statistical Methods in Education and Psychology**. MA: Allyn & Bacon.

Olkin, Ingram., & Finn, Jeremy. (1990). Testing correlated correlations. **Psychological Bulletin** 108 2: 330-333.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. **Psychological Bulletin** 87: 245-251.

Williams, E.J. (1959). The Comparison of regression variables. **Journal of the Royal Statistical Society, Series B, 21**, 396-399. quoted in Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. **Psychological Bulletin,** 87, 245-251.

ภาคผนวก

ต่อไปจะได้แสดงตัวอย่างของการเปรียบเทียบประสิทธิภาพของสูตรที่ใช้ในการทดสอบสมมุติฐานเปรียบเทียบค่าสหสัมพันธ์สองค่าที่มีตัวแปรเกณฑ์ตัวเดียวกันซึ่งมีสองสูตรคือของ Hotelling ที่เป็นสูตรเดิมและสูตรของ Meng, Rosenthal, & Rubin (1992) เพื่อให้เห็นข้อด้อยของอีกสูตรซึ่งจุดนี้น่าจะได้มีการจำลองเพื่อเปรียบเทียบต่อไป

ให้ สร้างตัวแปรสุ่ม X_1 และ X_2 โดยที่ทั้งสองตัวแปรมี ค่าเฉลี่ยเท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 และมีการแจกแจงปกติ โดยมีตัวแปรทั้งสองตั้งฉากกัน (orthogonal) กัน คือ ไม่มีความสัมพันธ์กันในเชิงเส้น จำนวนอย่างละ 99999 (N=99999) หลังจากนั้นสร้างตัวแปร y โดยที่ $y = (0.5)^{0.5}(X_1 + X_2)$ ซึ่งก็จะได้ตัวแปรสุ่ม y ที่มีค่าสหสัมพันธ์กับ X_1 และ X_2 เท่ากัน หลังจากนั้นได้ลองสุ่มเลือกข้อมูลมา 30 ชุด (n=30) โดยใช้คอมพิวเตอร์สุ่มและได้คำนวณค่าสถิติต่าง ๆ ดังนี้

$r_{yx1} = 0.81$

$r_{yx2} = 0.631$

$r_{x1x2} = 0.121$

หลังจากนั้นนำค่าสถิติเหล่านี้ไปแทนในสูตรจะได้ดังนี้

สูตรของ Hotelling ได้ค่า t = 2.9749 ซึ่งไม่มากกว่าค่า t วิกฤติที่ df = 27 ซึ่งมากกว่า 2.052 จึงปฏิเสธสมมุติฐานว่าง ทั้ง ๆ ที่จริง ๆ แล้วต้องไม่สามารถปฏิเสธสมมุติฐานว่างเนื่องจากค่าพารามิเตอร์ของประชากรเท่ากันคือมีค่าเท่ากับรากที่สองของ 0.5 ทั้งสองตัว

สูตรของ Meng, Rosenthal, & Rubin ได้ค่า z = 1.448 ซึ่งไม่สามารถปฏิเสธสมมุติฐานว่างได้เนื่องจากได้ค่า $z < 1.96$

จะเห็นได้ว่าสูตรของ Hotelling มีข้อด้อยกว่าและนำไปสู่การตัดสินใจเกี่ยวกับค่าพารามิเตอร์ที่ผิด แต่ทั้งนี้เป็นเพียงการยกตัวอย่างเพียงกรณีเดียวเท่านั้น ดังนั้นจึงต้องมีการทำซ้ำอีกหลาย ๆ ครั้งเพื่อความมั่นใจ