

๒.๒ วิธีการวิเคราะห์ข้อมูลโดยใช้สถิติเชิงอ้างอิง (Inference Statistics)

สำหรับสถิติที่ใช้ในการวิเคราะห์ข้อมูลเชิงอ้างอิง (Inference Statistics) หรือเรียกอีกอย่างหนึ่งว่า สถิติเชิงอนุมาน เป็นสถิติที่ใช้ผลการวิเคราะห์ข้อมูลที่ได้จากกลุ่มตัวอย่าง (Sample) หรือประชากรเพียงบางส่วน เพื่อนำมาอ้างอิงถึงประชากร (Population) อันเป็นการใช้ผลการวิเคราะห์ที่ได้จากข้อมูลที่มีอยู่จำกัดไปประมาณ อนุมาน บรรยาย หรือทำนายลักษณะบางประการของประชากรทั้งหมด โดยนำทฤษฎีความน่าจะเป็นมาประยุกต์ใช้

ดังนั้นการเลือกใช้หลักสถิติวิเคราะห์ข้อมูล จะต้องทำความเข้าใจเกี่ยวกับสัญลักษณ์ที่เกี่ยวข้อง ดังนี้

ค่าสถิติ	ค่าที่วัดได้จากประชากร (Parameter)	ค่าที่วัดได้จากตัวอย่าง (Sample)
ค่าเฉลี่ย	μ	$\bar{X}$
ส่วนเบี่ยงเบนมาตรฐาน	σ	S.D.
ความแปรปรวน	σ^2	S^2

ขั้นตอนการวิเคราะห์ข้อมูลโดยใช้สถิติอ้างอิง

ในการวิเคราะห์ข้อมูลโดยใช้สถิติอ้างอิงนั้น ผู้วิจัยจะต้องใช้วิธีการวิเคราะห์ข้อมูลตามวิธีการทางวิทยาศาสตร์ คือ

๑) **สังเกตเพื่อระบุปัญหา** หมายความว่า ผู้วิจัยต้องระบุถึงปัญหาการวิจัย มีปัญหาอะไร สำหรับปัญหาการวิจัยในการวิเคราะห์ด้วยสถิติอ้างอิงนี้ ผู้วิจัยได้ระบุไว้แล้วในปัญหาการวิจัย เช่น หัวข้อวิจัยเรื่อง “ความเห็นของบุคลากรต่องบบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย” ผู้วิจัยได้เขียนปัญหาการวิจัยไว้ว่า “บุคลากรมีความคิดเห็นต่องบบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตยแตกต่างกันหรือไม่?” โดยเป็นการกล่าวในภาพรวม

๒) **ขั้นตั้งสมมติฐาน** หมายความว่า เมื่อผู้วิจัยได้ระบุปัญหาแล้ว ผู้วิจัยจะต้องเริ่มตั้งสมมติฐานอันเป็นการหาคำตอบไว้ล่วงหน้า ซึ่งในงานวิจัยดังกล่าว ผู้วิจัยอาจตั้งสมมติฐานโดยจำแนกตาม เช่น

บุคลากรที่มีเพศต่างกัน มีความคิดเห็นต่องบบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกัน

บุคลากรที่มีอายุต่างกัน มีความคิดเห็นต่อบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกัน

บุคลากรที่มีประสบการณ์ต่างกัน มีความคิดเห็นต่อบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกัน

๓) ขั้นการรวบรวมข้อมูล หมายถึง ผู้วิจัยต้องทำการรวบรวมข้อมูล แล้วนำข้อมูลนั้นมาวิเคราะห์เพื่อตรวจสอบสมมติฐานที่ตั้งไว้ว่าถูกหรือผิด โดยมีหลักฐานยืนยัน ซึ่งในที่นี้ผู้วิจัยจะใช้สถิติเชิงอนุมานเข้ามาคำนวณเพื่อนำผลการวิเคราะห์ข้อมูลที่ได้นั้นนำไปเปรียบเทียบกับสมมติฐานที่ตั้งไว้ เช่น ผู้วิจัยต้องนำข้อมูลที่ได้จากการตอบแบบสอบถามของบุคลากรนั้นมาวิเคราะห์และจำแนกตามหลักการทางสถิติที่กำหนดไว้ เพื่อทำการพิสูจน์ว่า เป็นไปตามสมมติฐานที่ตั้งไว้หรือไม่?

๔) ขั้นสรุปผล หมายถึง เป็นขั้นตอนการสรุปว่าจะปฏิเสธ หรือยอมรับสมมติฐาน ตามหลักเหตุและผลเพื่อให้ได้คำตอบของปัญหา เช่น หากพบว่า บุคลากรทั้งชายและหญิง มีความเห็นต่อบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย เป็นไปในทิศทางเดียวกัน ก็ถือว่า เป็นการปฏิเสธสมมติฐานที่ตั้งไว้ แต่ถ้าบุคลากรทั้งชายและหญิง มีความเห็นต่อบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย ที่แตกต่างกัน ก็ถือว่า เป็นการยอมรับสมมติฐานที่ตั้งไว้ เป็นต้น

ดังนั้น ในการวิเคราะห์ข้อมูล จึงมีการดำเนินการ ดังนี้

๕.๑ ตั้งสมมติฐาน

๕.๒ กำหนดระดับนัยสำคัญ (Level of Significance)

๕.๓ หาค่าวิกฤต (Critical Value : CV)^๔

๕.๔ คำนวณสถิติโดยเลือกสถิติที่เหมาะสมในการทดสอบสมมติฐาน

๕.๕ พิจารณาผลการทดสอบสมมติฐาน

สำหรับสถิติที่ใช้ในการวิเคราะห์ข้อมูลเชิงอ้างอิงนั้น ประกอบด้วย

๑) สถิติวิเคราะห์ค่าที่ในกรณีทีกลุ่มตัวอย่างเป็นอิสระต่อกัน

(Independent Sample t-test)

๒) สถิติวิเคราะห์ค่าที่ในกรณีทีกลุ่มตัวอย่างสัมพันธ์กัน

(Paired Sample t-test)

^๔ โปรดอ่าน หลักในการทดสอบสมมติฐาน ในบทที่ ๖

- ๓) สถิติวิเคราะห์ความแปรปรวนทางเดียว (One Way ANOVA)
- ๔) สถิติวิเคราะห์ไค-สแควร์ (Chi-Square Test)
- ๕) สถิติวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's r)

๑) สถิติวิเคราะห์ค่าทีในกรณีที่กลุ่มตัวอย่างเป็นอิสระต่อกัน (Independent Sample t-test)

เป็นสถิติวิเคราะห์เพื่อทำการเปรียบเทียบตัวแปร ซึ่งมีเงื่อนไขดังนี้

๑.๑ เปรียบเทียบเฉพาะตัวแปรย่อย ๒ ตัวแปรเท่านั้น เช่น เปรียบเทียบความเห็นของชายกับหญิง, เปรียบเทียบความเห็นของนักเรียนในโรงเรียน ก. กับ โรงเรียน ข. เป็นต้น

๑.๒ ตัวแปรต้น เป็นตัวแปรกลุ่ม (Nominal Scale) หรือ ตัวแปรเรียงลำดับ (Ordinal Scale) และตัวแปรตามเป็นตัวแปรอันตรภาคชั้น (Interval Scale) หรือ ตัวแปรอัตราส่วน (Ratio Scale)

๑.๓ ใช้ค่าสถิติแบบมี Parameter กล่าวคือ ใช้ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐานมาใช้เป็นฐานในการคำนวณ

ยกตัวอย่างเช่น

พนักงานองค์การบริหารส่วนตำบล ๑๐๐ คน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย ในเรื่องการส่งเสริมให้ประชาชนได้ไปใช้สิทธิเลือกตั้ง โดยเฉลี่ยเท่ากับ ๓.๖ ส่วนเบี่ยงเบนมาตรฐานอยู่ที่ ๑.๓๕ เมื่อจำแนกตามเพศ มีค่าเฉลี่ยดังนี้

เพศ	จำนวน	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน
ชาย	df1	๔.๐๐	๑.๐๐๐
หญิง	df2	๓.๒๐	๑.๕๖๕
รวม	๑๐๐	๓.๖๐	๑.๓๕๐

จากข้อมูลดังกล่าว ผู้วิจัยจึงต้องการศึกษาวิเคราะห์ว่า พนักงานองค์การบริหารส่วนตำบลที่มีเพศต่างกัน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกันหรือไม่

ดังนั้น ในการวิเคราะห์ข้อมูล จึงมีการดำเนินการดังนี้

๑) ตั้งสมมติฐาน

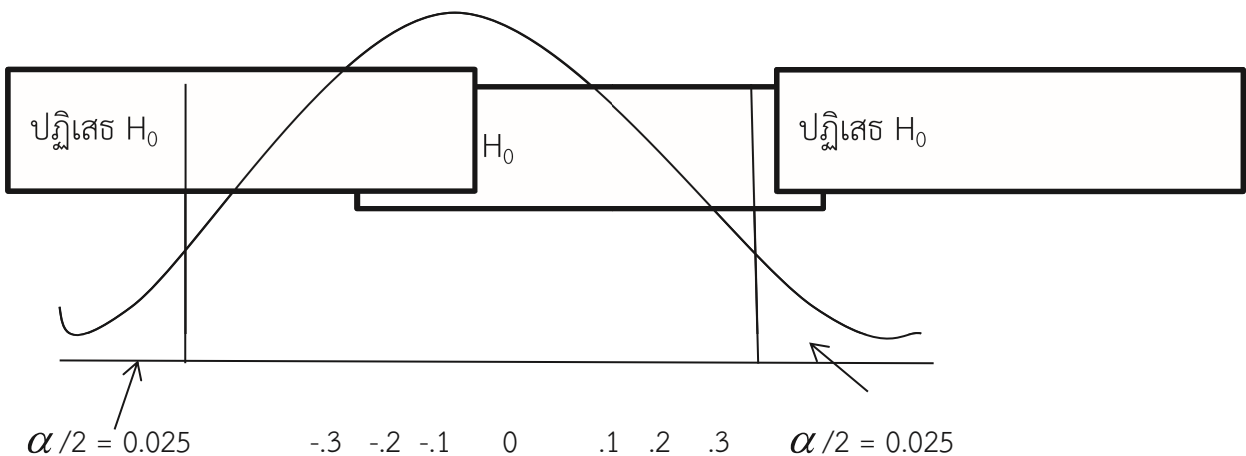
สมมติฐานทางสถิติ : สมมติฐานการวิจัย

$H_0 : \mu_1 = \mu_2$: บุคลากรที่มีเพศต่างกัน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย ไม่แตกต่างกัน

$H_1 : \mu_1 \neq \mu_2$: บุคลากรที่มีเพศต่างกัน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกัน

๒) กำหนดระดับนัยสำคัญ (Level of Significance)

กำหนดนัยสำคัญทางสถิติที่ ๐.๐๕ ซึ่งสามารถนำมาเขียนเป็นภาพได้ดังนี้



ภาพที่ ๓๕ แสดงการกำหนดนัยสำคัญทางสถิติที่ ๐.๐๕

๓) หาค่าวิกฤต (Critical Value : CV)

ผู้วิจัยหาค่าสถิติที่ใช้เป็นจุดแบ่งระหว่างบริเวณการยอมรับและบริเวณการปฏิเสธสมมติฐาน โดยการเปิดตารางค่าวิกฤตได้ จากหนังสือสถิติทั่วไป ภาคผนวก ๑ โดยดูจากค่าชั้นของความอิสระ (Degree of Freedom) ซึ่งสถิติวิเคราะห์ค่าที่ (t-test) มี ๒ แบบ คือ ๑) แบบที่กรณีประชากรทั้ง ๒ กลุ่มมีความแปรปรวนของข้อมูลไม่แตกต่างกัน และ ๒) แบบที่กรณีที่ประชากรทั้งสองกลุ่มมีความแปรปรวนของข้อมูลแตกต่างกัน

ดังนั้น ก่อนที่จะคำนวณหาค่า t-test เพื่อทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากรทั้ง ๒ กลุ่ม จะต้องพิจารณาว่าประชากรทั้ง ๒ กลุ่มมีค่าความแปรปรวนของข้อมูลต่างกันหรือไม่ โดยมีการตั้งสมมติฐานการวิจัย

$H_0 : \mu_1 = \mu_2$: มีความแปรปรวน ไม่แตกต่างกัน

$H_1 : \mu_1 \neq \mu_2$: มีความแปรปรวน แตกต่างกัน

จากนั้น ผู้วิจัยจึงทำการวิเคราะห์ความแตกต่างของข้อมูล โดยนำ

ค่าความแปรปรวนของเพศชาย **หารด้วย** ค่าความแปรปรวนของเพศหญิง ดังนี้

$$F = \frac{1.000}{2.249} = 0.44464$$

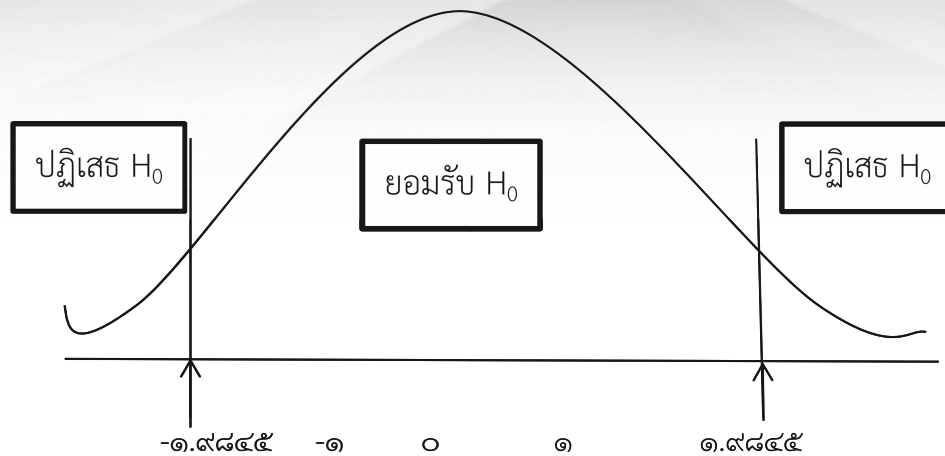
ดังนั้น ค่า F คำนวณ จึงเท่ากับ 0.44464 เมื่อเปิดตาราง F ที่ช่อง p50 w₁ และ ช่อง df2 (ภาคผนวกหน้า ๒๙๐) ได้ค่า f ตาราง เท่ากับ ๑.๖๐ ซึ่งค่า F ที่คิดได้น้อยกว่า F ตาราง จึงปฏิเสธคือ แสดงว่า ค่าความแปรปรวนไม่แตกต่างกัน จึงนำไปทดสอบในกรณีมีค่าความแปรปรวนไม่แตกต่างกัน

เมื่อผู้วิจัยพบว่า ความแปรปรวนของข้อมูล ๒ ชุดนี้ ไม่แตกต่างกันแล้ว ผู้วิจัยทำการหาช่วงวิกฤตของค่า t โดยดูจากค่าชั้นของความอิสระ (Degree of Freedom)

$$\text{โดยที่ } df = n_1 + n_2 - 2$$

$$\text{แทนค่า } df = 50 + 50 - 2 = 98$$

เมื่อเปิดตารางในภาคผนวก หน้า ๒๘๗ ที่ df. = 98 ณ ระดับนัยสำคัญที่ 0.05 (Two tail) ได้ค่า t_{ตาราง} = ๑.๙๘๔๕ หมายความว่า ค่า t_{คำนวณ} มากกว่า t_{ตาราง} จึงยอมรับสมมติฐาน ดังภาพ



ภาพที่ ๓๖ แสดงการกำหนดค่านัยสำคัญทางสถิติที่ ๐.๐๕ จากการเปิดตาราง

จากภาพที่ ๓๕ จะพบว่า จุดตัดของค่าวิกฤตที่ ๐.๐๕ คือ ค่าของ $t_{ตาราง}$ ที่ ๑.๙๘๔๕ หมายความว่า หากค่า $t_{คำนวณ}$ มากกว่า $t_{ตาราง}$ จึงยอมรับสมมติฐาน กล่าวคือ พนักงานองค์การบริหารส่วนตำบลที่มีเพศต่างกันมีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกัน แต่ถ้าหากค่า $t_{คำนวณ}$ น้อยกว่า $t_{ตาราง}$ จึงปฏิเสธสมมติฐาน กล่าวคือ พนักงานองค์การบริหารส่วนตำบลที่มีเพศต่างกันมีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย ไม่แตกต่างกัน

๔) คำนวนโดยใช้สถิติ

สูตร

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

โดยที่

$$S^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

แทนค่าในสูตร

$$\begin{aligned}
 S^2 &= \frac{(50 - 1) 1.00^2 + (50 - 1) 1.56^2}{50 + 50 - 2} \\
 &= \frac{49 + 119.856}{98} \\
 &= 1.223
 \end{aligned}$$

เมื่อได้ ค่า S^2 แล้วจึงนำมาวิเคราะห์ด้วยสูตร t test ดังต่อไปนี้

$$\begin{aligned}
 t &= \frac{4.00 - 3.20}{\sqrt{1.223 \left[\frac{1}{50} + \frac{1}{50} \right]}} \\
 &= \frac{0.80}{\sqrt{1.223 \times 0.04}} \\
 &= \frac{0.80}{\sqrt{0.04892}} \\
 &= \frac{0.80}{0.22125} \\
 &= 3.617
 \end{aligned}$$

๕) พิจารณาผลการทดสอบสมมติฐาน

ผลการทดสอบสมมติฐาน พบว่า ค่า $t_{คำนวณ} = 3.617$ ซึ่งมากกว่า $t_{ตาราง} = 1.9845$ ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ ดังนั้น จึงสรุปได้ว่า บุคลากรที่เพศต่างกัน มีความคิดเห็นต่อบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกันอย่างมีนัยสำคัญ จึงเป็นไปตามสมมติฐานที่ตั้งไว้

๒) สถิติวิเคราะห์ค่าที่ในกรณีที่กลุ่มตัวอย่างสัมพันธ์กัน (Paired Sample t-test)

เป็นสถิติวิเคราะห์เพื่อทำการเปรียบเทียบตัวแปรที่มีความสัมพันธ์กัน หรือ เป็นการทดสอบจากประชากรเพียงกลุ่มเดียว แต่จะทำการทดสอบซ้ำ ๒ ครั้ง โดยจะทำการทดสอบแต่ละครั้งจะต้องมีระยะเวลาที่ห่างกัน^๕ ซึ่งมีเงื่อนไขดังนี้

๒.๑ เปรียบเทียบเฉพาะตัวแปรย่อย ๒ ตัวแปร ได้มาจากกลุ่มตัวอย่างกลุ่มเดียวกัน หรือมาจากกลุ่มตัวอย่าง ๒ กลุ่มที่มีความสัมพันธ์กัน

๒.๒ ค่าของตัวแปรตามในแต่ละหน่วยเป็นอิสระต่อกัน และการวัดค่าตัวแปรตามอยู่ในมาตราอันตรภาคชั้น หรืออัตราส่วน

๒.๓ กลุ่มตัวอย่างได้มาโดยการสุ่มจากประชากรที่มีการแจกแจงแบบปกติ

๒.๔ การแจกแจงของความแตกต่างระหว่างข้อมูล หรือคะแนนแต่ละคู่เป็นแบบปกติ

๒.๕ ไม่ทราบค่าความแปรปรวนของประชากร

^๕ สุรพล พรหมกุล, ระเบียบวิธีวิจัยทางรัฐศาสตร์, (กรุงเทพมหานคร: สำนักพิมพ์ โอเดียนสโตร์), ๒๕๕๔, หน้า ๑๘๐.

ยกตัวอย่างเช่น

นิสิตที่ศึกษาในรายวิชาการปกครองส่วนท้องถิ่นไทย ได้ทำการทดสอบความรู้เกี่ยวกับการปกครองท้องถิ่นของไทย ก่อนเรียน และหลังเรียน มีผลคะแนน ดังนี้

คนที่	ครั้งที่ ๑	ครั้งที่ ๒
๑	๒๒	๒๕
๒	๒๕	๒๗
๓	๒๔	๒๘
๔	๒๐	๒๗
๕	๒๒	๒๔
๖	๒๔	๒๕
๗	๒๓	๒๑
๘	๒๐	๒๔
๙	๒๙	๒๓
๑๐	๓๒	๓๖
๑๑	๒๖	๒๙
๑๒	๒๗	๒๙
๑๓	๓๐	๒๘
๑๔	๓๐	๓๑
๑๕	๒๘	๓๑

จากข้อมูลดังกล่าว ผู้วิจัยจึงต้องการศึกษาวิเคราะห์ว่า นิสิตมีผลการเรียนรู้เกี่ยวกับการปกครองท้องถิ่นของไทยหลังเรียนสูงกว่าก่อนเรียนหรือไม่

ดังนั้น ในการวิเคราะห์ข้อมูล จึงมีการดำเนินการดังนี้

๑) ตั้งสมมติฐาน

สมมติฐานทางสถิติ : สมมติฐานทางการวิจัย

$H_0 : \mu_1 = \mu_2$: นิสิตมีผลการเรียนรู้เกี่ยวกับการปกครองท้องถิ่นของไทยหลังเรียน เท่ากับ ก่อนเรียน

$H_1 : \mu_1 > \mu_2$: นิสิตมีผลการเรียนรู้เกี่ยวกับการปกครองท้องถิ่นของ
ไทยหลังเรียน สูงกว่า ก่อนเรียน

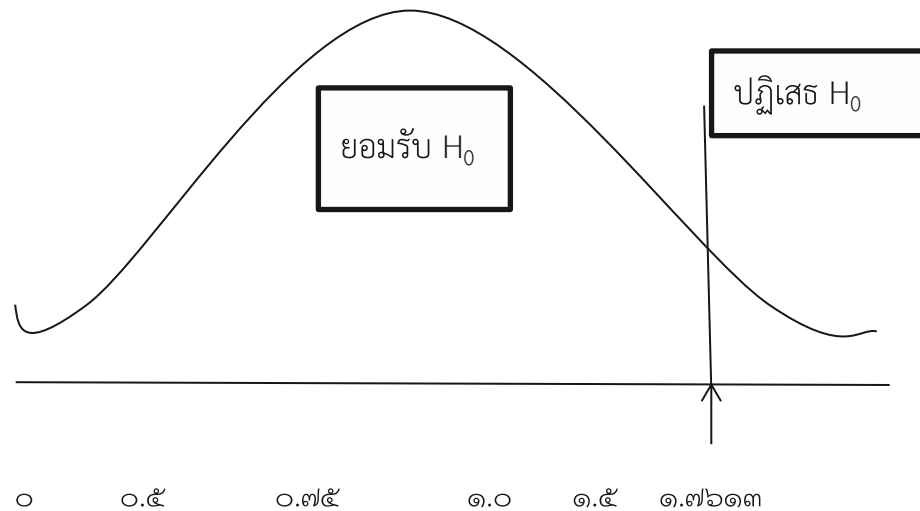
๒) กำหนดระดับนัยสำคัญ (Level of Significance)

กำหนดนัยสำคัญทางสถิติที่ ๐.๐๕

๓) หาค่าวิกฤต (Critical Value : CV)

ผู้วิจัยหาค่าสถิติที่ใช้เป็นจุดแบ่งระหว่างบริเวณการยอมรับและ
บริเวณการปฏิเสธสมมติฐาน โดยการเปิดตารางค่าวิกฤตได้ โดยที่ $df = n - 1$ แทนค่า df
 $= ๑๕ - ๑ = ๑๔$

เมื่อเปิดตารางในภาคผนวก หน้า ๒๘๕ ที่ $df = ๑๔$ ที่นัยสำคัญทางสถิติ $= ๐.๐๕$
ทดสอบแบบทางเดียว ได้ค่า $t_{ตาราง} = ๑.๗๖๑๓$ หมายความว่า ค่า $t_{คำนวณ}$ มากกว่า $t_{ตาราง}$ จึง
ยอมรับสมมติฐาน ดังภาพ



ภาพที่ ๓๗ แสดงการกำหนดค่านัยสำคัญทางสถิติที่ ๐.๐๕ จากการเปิดตาราง

จากภาพที่ ๓๗ จะพบว่า จุดตัดของค่าวิกฤตที่ ๐.๐๕ คือ ค่าของ $t_{ตาราง}$ ที่ ๑.๗๖๑๓
หมายความว่า หากค่า $t_{คำนวณ}$ มากกว่า $t_{ตาราง}$ จึงยอมรับสมมติฐาน กล่าวคือ นิสิตมีผลการ
เรียนรู้เกี่ยวกับการปกครองท้องถิ่นของไทยหลังเรียน สูงกว่า ก่อนเรียน แต่ถ้าหากค่า $t_{คำนวณ}$

น้อยกว่า $t_{ตาราง}$ จึงปฏิเสธสมมติฐาน กล่าวคือ นิสิตมีผลการเรียนรู้เกี่ยวกับการปกครองท้องถิ่นของไทยหลังเรียน ต่ำกว่าหรือเท่ากับ ก่อนเรียน

๔) คำนวณโดยใช้สถิติ

สูตร^{๑๐}

$$t = \frac{\sum D}{\sqrt{\frac{n \sum D^2 - (\sum D)^2}{n-1}}}$$

เมื่อ D = ผลต่างระหว่างข้อมูลแต่ละคู่

n = จำนวนกลุ่มตัวอย่าง

แทนค่าในสูตรดังตารางดังนี้

คนที่	ครั้งที่ ๑	ครั้งที่ ๒	D	D ^๒
๑	๒๒	๒๕	๓	๙
๒	๒๕	๒๗	๒	๔
๓	๒๔	๒๘	๔	๑๖
๔	๒๐	๒๗	๗	๔๙
๕	๒๒	๒๔	๒	๔
๖	๒๔	๒๕	๑	๑
๗	๒๓	๒๑	-๒	๔
๘	๒๐	๒๔	๔	๑๖
๙	๒๙	๒๓	-๖	๓๖
๑๐	๓๒	๓๖	๔	๑๖
๑๑	๒๖	๒๙	๓	๙
๑๒	๒๗	๒๙	๒	๔
๑๓	๓๐	๒๘	-๒	๔
๑๔	๓๐	๓๑	๑	๑
๑๕	๒๘	๓๑	๓	๙
รวม			๒๖	๑๘๒

^{๑๐} วาโร เฟ็งสวัสต์, ผศ.ดร., วิธีวิทยาการวิจัย, (กรุงเทพฯ: สุวีริยาสาส์น), ๒๕๕๑, หน้า ๓๓๙ -

นำมาวิเคราะห์ด้วยสูตร t test ดังต่อไปนี้

$$\begin{aligned}
 t &= \frac{26}{\sqrt{\frac{15(182) - 26^2}{15 - 1}}} \\
 &= \frac{26}{\sqrt{\frac{2730 - 676}{14}}} \\
 &= \frac{26}{\sqrt{146.7}} \\
 &= \frac{26}{12.11} \\
 &= 2.147
 \end{aligned}$$

๕) พิจารณาผลการทดสอบสมมติฐาน

ผลการทดสอบสมมติฐาน พบว่า ค่า $t_{คำนวณ} = 2.147$ ซึ่งมากกว่า $t_{ตาราง} = 1.7613$ ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ ดังนั้น จึงสรุปได้ว่า นิสิตมีผลการเรียนรู้เกี่ยวกับการปกครองท้องถิ่นของไทยหลังเรียนสูงกว่าก่อนเรียน อย่างมีนัยสำคัญทางสถิติ จึงเป็นไปตามสมมติฐานที่ตั้งไว้

๓) สถิติวิเคราะห์ความแปรปรวนทางเดียว (One Way ANOVA)

การวิเคราะห์ความแปรปรวนทางเดียว เป็นการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของตัวแปรอิสระเพียงตัวเดียว แต่จำแนกออกเป็น ๒ ตัวแปรขึ้นไป เป็นสถิติวิเคราะห์เพื่อทำการเปรียบเทียบตัวแปร ซึ่งมีเงื่อนไขดังนี้

๑.๑ เปรียบเทียบตัวแปรย่อยตั้งแต่ ๒ ตัวแปรขึ้นไป

๑.๒ ตัวแปรต้น เป็นตัวแปรกลุ่ม (Nominal Scale) หรือ ตัวแปรเรียงลำดับ (Ordinal Scale) และตัวแปรตามเป็นตัวแปรอันตรภาคชั้น (Interval Scale) หรือ ตัวแปรอัตราส่วน (Ratio Scale)

๑.๓ ใช้ค่าสถิติแบบมี Parameter

ข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนทางเดียว

๑. กลุ่มตัวอย่างได้มาโดยวิธีการสุ่ม และเป็นอิสระจากกัน ซึ่งได้มาจากประชากรที่มีการแจกแจงแบบปกติ
๒. ความแปรปรวนของประชากรแต่ละกลุ่มต้องเท่ากัน
๓. ตัวแปรอิสระมีเพียงตัวแปรเดียว แต่จำแนกตั้งแต่ ๒ ระดับขึ้นไป
๔. ตัวแปรตามมีเพียงตัวเดียว โดยผลที่วัดได้จากตัวแปรตามนี้เป็นตัวแปรอันตรภาคชั้น (Interval Scale) หรือ ตัวแปรอัตราส่วน (Ratio Scale)

หลักการวิเคราะห์ความแปรปรวนทางเดียว

การวิเคราะห์ความแปรปรวนทางเดียว จะมีการคำนวณหาค่าความแปรปรวน ซึ่งค่าความแปรปรวนนี้เกิดจาก ๒ แหล่ง คือ ความแปรปรวนระหว่างกลุ่ม และความแปรปรวนภายในกลุ่ม

๑) ความแปรปรวนระหว่างกลุ่ม (Between-group variance) เป็นค่าที่แสดงให้เห็นถึงขนาดของความแตกต่างระหว่างค่าเฉลี่ยของกลุ่มต่างๆ ซึ่งเป็นผลสำคัญของตัวแปรอิสระที่ศึกษา

๒) ความแปรปรวนภายในกลุ่ม (Within-group variance) เป็นค่าที่แสดงให้เห็นถึงคะแนนแต่ละตัวที่รวบรวมได้ภายในกลุ่ม ซึ่งเป็นผลของตัวแปรอื่นๆ ที่ทำให้คลาดเคลื่อนไป ค่าที่คำนวณได้นี้เรียกว่า ค่าความคลาดเคลื่อน

ดังนั้น จึงสามารถนำมาเขียนเป็นสูตร ได้ดังนี้

$$F = \frac{MS_b}{MS_w}$$

โดยที่

$$MS_b = \frac{SS_b}{K - 1}$$

โดยที่

$$SS_b = \sum n_j (\bar{X}_j - \bar{X})^2$$

$$MS_w = \frac{SS_w}{n - k}$$

โดยที่

$$SS_w = SS_t - SS_b$$

$$\text{ซึ่ง } SS_t = \sum \sum (X_{ij} - \bar{X})^2$$

เมื่อ

F	=	อัตราส่วนของความแปรปรวน
MS _b	=	ค่าความแปรปรวนระหว่างกลุ่ม
MS _w	=	ค่าความแปรปรวนภายในกลุ่ม
SS _b	=	ผลรวมของกำลังสองของค่าเบี่ยงเบนมาตรฐานระหว่างกลุ่ม
SS _w	=	ผลรวมของกำลังสองของค่าเบี่ยงเบนมาตรฐานภายในกลุ่ม
SS _t	=	ผลรวมของกำลังสองของค่าเบี่ยงเบนทั้งหมด
K	=	จำนวนกลุ่ม
n	=	จำนวนตัวอย่างทั้งหมด
n _j	=	จำนวนตัวอย่างของกลุ่ม j
$\bar{X}_j$	=	ค่าเฉลี่ยของกลุ่ม j
$\bar{X}_t$	=	ค่าเฉลี่ยทั้งหมด
df _b	=	ช่วงชั้นความเป็นอิสระภายในกลุ่ม
df _w	=	ช่วงชั้นความเป็นอิสระระหว่างกลุ่ม ^{๑๑}

^{๑๑} อ้างแล้ว หน้า ๓๔๓ - ๓๔๗.

ยกตัวอย่างเช่น

พนักงานองค์การบริหารส่วนตำบล ๑๕ คน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตยในเรื่องการส่งเสริมให้ประชาชนได้ไปใช้สิทธิเลือกตั้ง เมื่อจำแนกตามอายุ มีดังนี้

	คะแนนของแต่ละช่วงอายุ			รวม
	ต่ำกว่า ๓๐ ปี	๓๐ - ๔๙ ปี	๕๐ ปีขึ้นไป	
คนที่ ๑	๕ (๒๕)	๓ (๙)	๓ (๙)	
คนที่ ๒	๕ (๒๕)	๔ (๑๖)	๒ (๘)	
คนที่ ๓	๕ (๒๕)	๕ (๒๕)	๓ (๙)	
คนที่ ๔	๔ (๑๖)	๓ (๙)	๒ (๘)	
คนที่ ๕	๓ (๙)	๒ (๘)	๑ (๑)	
รวมคะแนน T	๒๒	๑๗	๑๑	๕๐
ค่าเฉลี่ย ($\bar{X}$)	๔.๔	๓.๔	๒.๒	๓.๓๓
ผลรวมของคะแนนยกกำลังสอง ($\sum X^2$)	๑๐๐	๖๓	๒๗	๑๙๐

$N = ๑๕$

จากข้อมูลดังกล่าว ผู้วิจัยจึงต้องการศึกษาวิเคราะห์ว่า พนักงานองค์การบริหารส่วนตำบลที่มีอายุต่างกัน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกันหรือไม่

วิธีทำ

๑) ตั้งสมมติฐาน

$H_0 : \mu_1 = \mu_2 = \mu_3$: บุคลากรที่มีอายุต่างกัน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย ไม่แตกต่างกัน

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$: บุคลากรที่มีอายุต่างกัน มีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกัน

๒) กำหนดระดับนัยสำคัญ (Level of Significance)

กำหนดนัยสำคัญทางสถิติที่ ๐.๐๕

๓) หาค่าวิกฤต (Critical Value : CV)

ผู้วิจัยหาค่าสถิติที่ใช้เป็นจุดแบ่งระหว่างบริเวณการยอมรับและบริเวณการปฏิเสธสมมติฐาน โดยการเปิดตารางค่าวิกฤตได้ ตารางสถิติใน ภาคผนวก โดยดูจากค่าชั้นของความอิสระ (Degree of Freedom) เมื่อเปิดตาราง F ในภาคผนวกหน้า ๒๙๐ ที่ช่อง df1 ซึ่งหาได้จากค่า df_b (ช่วงชั้นความเป็นอิสระภายในกลุ่ม) $= k - 1 = 3 - 1 = 2$ และ ช่อง df2 ซึ่งหาได้จากค่า df_w (ช่วงชั้นความเป็นอิสระระหว่างกลุ่ม) $= n - k = 100 - 3 = 97$ ได้ค่าเท่ากับ ๓.๐๙ ดังนั้น ค่า $F_{ตาราง} = 3.09$ หมายความว่า ค่า $F_{คำนวณ}$ มากกว่า $F_{ตาราง}$ จึงยอมรับสมมติฐาน กล่าวคือ พนักงานองค์การบริหารส่วนตำบลที่มีอายุต่างกันมีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย แตกต่างกัน แต่ถ้าหากค่า $F_{คำนวณ}$ น้อยกว่า $F_{ตาราง}$ จึงปฏิเสธสมมติฐาน กล่าวคือ พนักงานองค์การบริหารส่วนตำบลที่มีเพศต่างกันมีความคิดเห็นเกี่ยวกับบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตย ไม่แตกต่างกัน

๔) คำนวณโดยใช้สถิติ

จากข้อมูล แทนค่าได้ดังนี้

$$\begin{aligned}
 K (\text{จำนวนกลุ่ม}) &= 3 \\
 n (\text{จำนวนตัวอย่างทั้งหมด}) &= 15 \\
 n_j (\text{จำนวนตัวอย่างของกลุ่ม } j) &= n_1 = 5, n_2 = 5, n_3 = 5 \\
 \bar{X}_j (\text{ค่าเฉลี่ยของกลุ่ม } j) &= \bar{X}_1 = 4.4, \bar{X}_2 = 3.4, \bar{X}_3 = 2.2 \\
 \bar{X}_t (\text{ค่าเฉลี่ยทั้งหมด}) &= \bar{X}_t = 3.33 \\
 df_b (\text{ช่วงชั้นความเป็นอิสระภายในกลุ่ม}) &= k - 1 = 3 - 1 = 2 \\
 df_w (\text{ช่วงชั้นความเป็นอิสระระหว่างกลุ่ม}) &= n - k = 15 - 3 = 12
 \end{aligned}$$

๑. หาค่า SS_t (ผลรวมของกำลังสองของค่าเบี่ยงเบนทั้งหมด)

สูตร

$$SS_t = \sum \sum (X_{ij} - \bar{X}_t)^2$$

แทนค่า

$$\begin{aligned}
 SS_t &= (5 - 3.33)^2 + (5 - 3.33)^2 + (5 - 3.33)^2 + (4 - 3.33)^2 + \\
 &\quad (3 - 3.33)^2 + (3 - 3.33)^2 + (4 - 3.33)^2 + (5 - 3.33)^2 + \\
 &\quad (3 - 3.33)^2 + (2 - 3.33)^2 + (3 - 3.33)^2 + (2 - 3.33)^2 +
 \end{aligned}$$

$$\begin{aligned}
 & (3 - 3.33)^2 + (2 - 3.33)^2 + (1 - 3.33)^2 \\
 &= 2.7789 + 2.7789 + 2.7789 + 0.4444 + 0.1111 + \\
 & 0.1111 + 0.4444 + 2.7789 + 0.1111 + 1.7778 + \\
 & 0.1111 + 1.7778 + 0.1111 + 1.7778 + 5.5556 \\
 &= 23.3333
 \end{aligned}$$

๒. หาค่า SS_b (ผลรวมของกำลังสองของค่าเบี่ยงเบนมาตรฐานระหว่างกลุ่ม)

สูตร

$$SS_b = \sum n_j (\bar{X}_j - \bar{X}_t)^2$$

แทนค่า

$$\begin{aligned}
 SS_b &= 5 (4.4 - 3.33)^2 + 5 (3.4 - 3.33)^2 + 5 (2.2 - 3.33)^2 \\
 &= 5.7245 + 0.2045 + 1.2765 \\
 &= 7.2055
 \end{aligned}$$

๓. หาค่า SS_w (ผลรวมของกำลังสองของค่าเบี่ยงเบนมาตรฐานภายในกลุ่ม)

สูตร

$$SS_w = SS_t - SS_b$$

แทนค่า

$$\begin{aligned}
 SS_w &= 23.3333 - 7.2055 \\
 &= 16.1278
 \end{aligned}$$

๔. หาค่า MS_b (ค่าความแปรปรวนระหว่างกลุ่ม)

สูตร

$$MS_b = \frac{SS_b}{K-1}$$

แทนค่า

$$\begin{aligned}
 MS_b &= 7.2055 / 3 - 1 = 2 \\
 &= 3.60275
 \end{aligned}$$

๕. หาค่า MS_w (ค่าความแปรปรวนภายในกลุ่ม)

สูตร

$$MS_w = \frac{SS_w}{n-k}$$

แทนค่า

$$MS_w = 16.1276/15 - 3 = 12$$

$$= 1.34347$$

๖. หาค่า F (อัตราส่วนของความแปรปรวน)

สูตร

$$F = \frac{MS_b}{MS_w}$$

แทนค่า

$$F = 3.60255/1.34347$$

$$= 2.661$$

นำข้อมูลมาบรรจุลงในตาราง จะได้ดังนี้

ความแปรปรวน	SS	Df	MS	F
ระหว่างกลุ่ม (b)	๗.๒๐๕๙	๒	๓.๖๐๒๕๕	๒.๖๖๑
ภายในกลุ่ม (w)	16.1276	12	1.34347	
รวม	๒๓.๓๓๓๕	๑๔		

๕) พิจารณาผลการทดสอบสมมติฐาน

ผลการทดสอบสมมติฐาน พบว่า ค่า $F_{คำนวณ} = 2.661$ ซึ่งน้อยกว่า $F_{ตาราง} = 3.05$ ซึ่งไม่เป็นไปตามสมมติฐานที่ตั้งไว้ ดังนั้น จึงสรุปได้ว่า บุคลากรที่อายุต่างกัน มีความคิดเห็นต่อบทบาทขององค์การบริหารส่วนตำบลกับการส่งเสริมประชาธิปไตยไม่แตกต่างกัน จึงไม่เป็นไปตามสมมติฐานที่ตั้งไว้

๔) สถิติวิเคราะห์ไคแอสควร์ (Chi-Square Test)

สถิติวิเคราะห์ไคแอสควร์ (Chi-Square Test : χ^2) คือ อัตราส่วนระหว่างกำลังสองของผลต่างของความถี่ที่ได้จากการสังเกต (Observed frequency : O) กับความถี่ที่คาดหวัง (Expected frequency : E) ตามสมมติฐานที่ตั้งขึ้น โดยทำการหาค่าความสัมพันธ์ระหว่างค่าความถี่ที่สังเกตได้กับค่าความถี่ที่คาดหวัง กล่าวคือ หากค่าความถี่ที่สังเกตได้กับค่าความถี่ที่คาดหวังมีความแตกต่างกันมากเท่าใด แสดงว่า มีความสัมพันธ์กันมากเท่านั้น กล่าวคือ มีความแตกต่างกันนั่นเอง และสถิตินี้สามารถใช้ทดสอบได้กับข้อมูลทุกชนิด แต่ส่วนมากจะใช้กับข้อมูลในรูปของความถี่ หรือข้อมูลที่จัดเป็นความถี่ได้

การคำนวณแบ่งออกได้เป็น 2 กรณี คือ กรณีที่มีตัวแปรเดียวกับกรณีที่มีสองตัวแปร

๑.กรณีตัวแปรเดียว หมายถึง ตัวแปรนั้นแบ่งเป็นลักษณะหรือกลุ่มย่อย ซึ่งต้องการทดสอบว่า จำนวนหรือสัดส่วนในกลุ่มย่อยๆ นั้นเท่ากันหรือไม่หรือเท่ากับสัดส่วนที่เราคาดหวังหรือไม่ คำนวณได้จากสูตร^{๑๒} ดังนี้

$$X^2 = \sum \frac{(O-E)^2}{E}$$

ค่า df = k - 1

เมื่อ k คือ จำนวนกลุ่ม

O คือ ค่าความถี่ที่สังเกตได้

E คือ ค่าความถี่ที่คาดหวัง

ตัวอย่างเช่น

การศึกษาความคิดเห็นของประชาชนต่อการปฏิรูปการเมือง โดยทำการศึกษาจากประชาชนใน ๔ ภาคของประเทศไทย จำนวน ๔๐๐ คน พบว่า มีประชาชนเห็นด้วยกับการปฏิรูปการเมือง จำนวน ๒๓๐ คน และไม่เห็นด้วย จำนวน ๑๗๐ คน จึงทดสอบว่าประชาชนเห็นด้วยกับการปฏิรูปการเมืองแตกต่างกันหรือไม่ ที่ระดับความเชื่อมั่น ๙๕%

๑) ตั้งสมมติฐาน

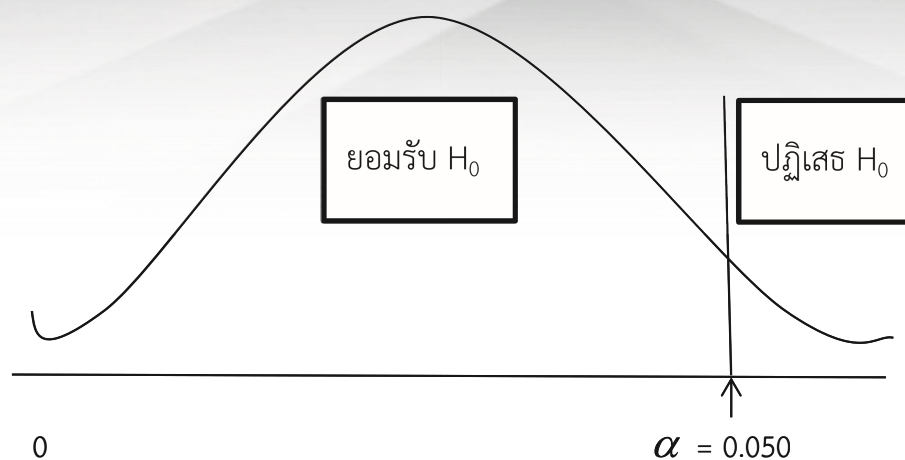
$H_0 : P_1 = P_2$: ประชาชนมีความคิดเห็นต่อการปฏิรูปการเมืองไม่แตกต่างกัน

$H_1 : P_1 \neq P_2$: ประชาชนมีความคิดเห็นต่อการปฏิรูปการเมือง แตกต่างกัน

๒) กำหนดระดับนัยสำคัญ (Level of Significance)

กำหนดนัยสำคัญทางสถิติที่ ๐.๐๕ ซึ่งสามารถนำมาเขียนเป็นภาพได้ ดังนี้

^{๑๒} วิเชียร ประเสริฐ และบัณฑิต บุณยสนอง, อ้างแล้ว, หน้า ๑๑๙.



ภาพที่ ๓๘ แสดงการกำหนดค่านัยสำคัญทางสถิติที่ ๐.๐๕

๓) หาค่าวิกฤต (Critical Value : CV)

ผู้วิจัยหาค่าสถิติที่ใช้เป็นจุดแบ่งระหว่างบริเวณการยอมรับและบริเวณการปฏิเสธสมมติฐาน โดยการเปิดตารางค่าวิกฤตได้จากหนังสือสถิติทั่วไป ภาคผนวก ๓ โดยดูจากค่าชั้นของความอิสระ (Degree of Freedom) เมื่อเปิดตารางไคแอสควร์ ในภาคผนวกหน้าที ๒๙๓ ที่ช่อง df ซึ่งหาได้จากค่า $k-1 = ๒ - ๑ = ๑$ ที่ช่องนัยสำคัญทางสถิติที่ ๐.๐๕ ได้ค่าเท่ากับ ๓.๘๔๑ ดังนั้น ค่า $X^2_{ตาราง} = ๓.๘๔๑$ หมายความว่า ค่า X^2 คำนวณมากกว่า $X^2_{ตาราง}$ จึงยอมรับสมมติฐาน กล่าวคือ ความคิดเห็นของประชาชนต่อการปฏิรูปการเมืองมีความสัมพันธ์กัน หมายความว่า มีความเห็นแตกต่างกัน แต่ถ้าหากค่า X^2 คำนวณน้อยกว่า $X^2_{ตาราง}$ จึงปฏิเสธสมมติฐาน กล่าวคือ ความคิดเห็นของประชาชนต่อการปฏิรูปการเมืองไม่มีความสัมพันธ์กัน กล่าวคือ มีความคิดเห็นไม่แตกต่างกัน

๔) คำนวณโดยใช้สถิติ

จากสูตรไคแอสควร์

$$X^2 = \sum \frac{(O-E)^2}{E}$$

$$\text{ค่า df} = k - 1$$

เมื่อ k คือ จำนวนกลุ่ม

O คือ ค่าความถี่ที่สังเกตได้

E คือ ค่าความถี่ที่คาดหวัง

ความคิดเห็น ของประชาชน ต่อการปฏิรูป การเมือง	ความถี่ที่ สังเกตได้ (O)	ความถี่ที่ คาดหวัง (E)	O - E	(O - E) ²
เห็นด้วย	๒๓๐	๒๐๐	๓๐	๙๐๐
ไม่เห็นด้วย	๑๗๐	๒๐๐	-๓๐	๙๐๐
รวม	๔๐๐	๔๐๐		

$$\begin{aligned}
 X^2 &= \frac{(230 - 200)^2}{200} + \frac{(170 - 200)^2}{200} \\
 &= 4.5 + 4.5 \\
 &= 9
 \end{aligned}$$

๕) พิจารณาผลการทดสอบสมมติฐาน

ผลการทดสอบสมมติฐาน พบว่า ค่า X^2 คำนวณ = ๙ ซึ่งมากกว่า X^2 ตาราง = ๓.๘๔๑๕ ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ ดังนั้น จึงสรุปได้ว่า มีความคิดเห็นของประชาชนต่อการปฏิรูปการเมืองไม่มีความสัมพันธ์กัน แตกต่างกัน จึงเป็นไปตามสมมติฐานที่ตั้งไว้

2. กรณีสองแปรขึ้นไป ในกรณีนี้จะดูความสัมพันธ์ระหว่างตัวแปรสองตัวขึ้นไป ซึ่งแต่ละตัวต่างเป็นตัวแปรระดับ Nominal Scale ตัวแปรแต่ละตัวแบ่งออกได้มากกว่าสองกลุ่ม สูตรที่ใช้ในการคำนวณเหมือนกันกับกรณีตัวแปรเดียว แต่วิธีการคำนวณค่าความถี่ที่คาดหวัง (Expected Frequency) โดยใช้หลักในการคำนวณได้จากสูตร ดังนี้

$$X^2 = \sum \frac{(O - E)^2}{E}$$

ค่า df = (r - 1)(c - 1)

เมื่อ r คือ จำนวนแถว

เมื่อ c คือ จำนวนคอลัมน์

$$\text{โดยที่ ค่า } E = \frac{R \times C}{N}$$

เมื่อ E คือ ความถี่คาดหวังที่ต้องการ

R คือ ผลรวมความถี่ของแถวที่ต้องการหาค่า E

C คือ ผลรวมความถี่ของคอลัมน์ที่ต้องการหาค่า E

N คือ จำนวนความถี่ทั้งหมด

ตัวอย่างเช่น

การศึกษาความคิดเห็นของประชาชนต่อการปฏิรูปการเมือง โดยทำการศึกษาจากประชาชนใน ๔ ภาคของประเทศไทย จำนวน ๔๐๐ คน โดยจำแนกเป็นรายภาคได้ดังนี้

ภูมิภาค	ความคิดเห็นของประชาชนต่อการปฏิรูปการเมือง		รวม
	เห็นด้วย	ไม่เห็นด้วย	
ภาคเหนือ	๔๐	๖๐	๑๐๐
ภาคอีสาน	๓๐	๗๐	๑๐๐
ภาคกลาง	๕๐	๕๐	๑๐๐
ภาคใต้	๘๐	๒๐	๑๐๐
รวม	๒๐๐	๒๐๐	๔๐๐

จากตารางดังกล่าว จงทดสอบว่าประชาชนในแต่ละภูมิภาคมีความเห็นต่อการปฏิรูปการเมืองแตกต่างกันหรือไม่ ที่ระดับความเชื่อมั่น ๙๕%

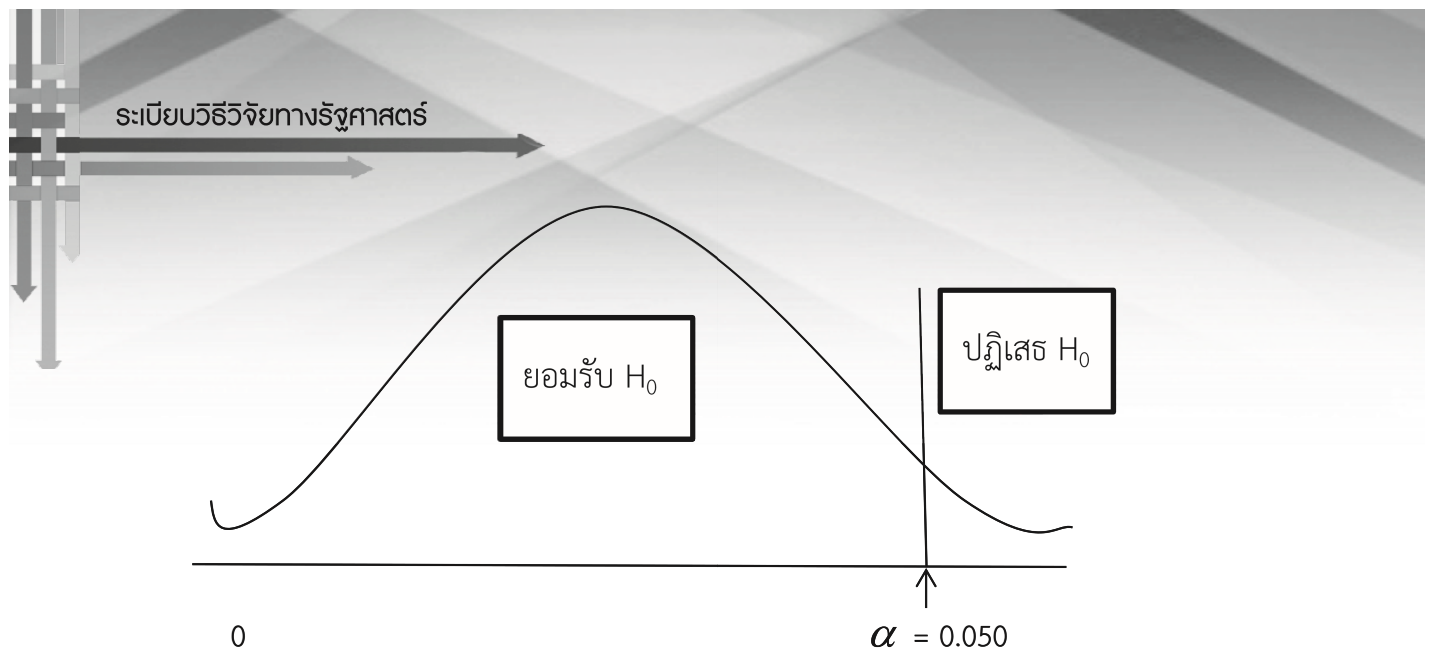
๑) ตั้งสมมติฐาน

$H_0 : \rho = 0$: ภูมิภาค **ไม่มีความสัมพันธ์**กับความคิดเห็นต่อการปฏิรูปการเมือง

$H_1 : \rho \neq 0$ ภูมิภาค **มีความสัมพันธ์**กับความคิดเห็นต่อการปฏิรูปการเมือง

๒) กำหนดระดับนัยสำคัญ (Level of Significance)

กำหนดนัยสำคัญทางสถิติที่ ๐.๐๕ ซึ่งสามารถนำมาเขียนเป็นภาพได้ดังนี้



ภาพที่ ๓๙ แสดงการกำหนดนัยสำคัญทางสถิติที่ ๐.๐๕

๓) หาค่าวิกฤต (Critical value : CV)

ผู้วิจัยหาค่าสถิติที่ใช้เป็นจุดแบ่งระหว่างบริเวณการยอมรับและบริเวณการปฏิเสธสมมติฐาน โดยการเปิดตารางค่าวิกฤตได้จากตารางสถิติในภาคผนวกหน้า ๒๙๓ โดยดูจากค่าชั้นของความอิสระ (Degree of Freedom) เมื่อเปิดตารางไคแอสควร์ที่ช่อง df ซึ่งหาได้จากจำนวนแถวลบด้วย ๑ คูณกับจำนวนคอลัมน์ลบด้วย ๑ ($(r - 1)(c - 1)$) ดังนั้น ในการวิจัยนี้ มีจำนวน ๔ แถว และมีจำนวน ๒ คอลัมน์ ค่า $df = (4 - 1)(2 - 1) = 3$ ที่ช่องนัยสำคัญทางสถิติที่ ๐.๐๕ ได้ค่าเท่ากับ ๗.๘๑๕ ดังนั้น ค่า $X^2_{ตาราง} = ๗.๘๑๕$ หมายความว่า ค่า $X^2_{คำนวณ}$ มากกว่า $X^2_{ตาราง}$ จึงยอมรับสมมติฐาน กล่าวคือ ภูมิภาค มีความสัมพันธ์กับความคิดเห็นต่อการปฏิรูปการเมือง หมายความว่า มีความเห็นแตกต่างกัน แต่ถ้าหากค่า $X^2_{คำนวณ}$ น้อยกว่า $X^2_{ตาราง}$ จึงปฏิเสธสมมติฐาน กล่าวคือ ภูมิภาค ไม่มีความสัมพันธ์กับความคิดเห็นต่อการปฏิรูปการเมือง กล่าวคือ มีความคิดเห็นไม่แตกต่างกัน

๔) คำนวณโดยใช้สถิติ จากสูตรไคแอสควร์

$$X^2 = \sum \frac{(O - E)^2}{E}$$

ค่า $df = (r - 1)(c - 1)$

เมื่อ r คือ จำนวนแถว

เมื่อ c คือ จำนวนคอลัมน์

$$\text{โดยที่ ค่า } E = \frac{R \times C}{N}$$

เมื่อ E คือ ความถี่คาดหวังที่ต้องการ

R คือ ผลรวมความถี่ของแถวที่ต้องการหาค่า E

C คือ ผลรวมความถี่ของคอลัมน์ที่ต้องการหาค่า E

N คือ จำนวนความถี่ทั้งหมด

วิธีคิด

๑) หาค่าความถี่คาดหวังที่ต้องการ (E)

ภูมิภาค	ความคิดเห็นของประชาชนต่อการปฏิรูปการเมือง		รวม
	เห็นด้วย	ไม่เห็นด้วย	
ภาคเหนือ	$100 \times 200 = 50$ ----- ๕๐๐	$100 \times 200 = 50$ ----- ๕๐๐	๑๐๐
ภาคอีสาน	$100 \times 200 = 50$ ----- ๕๐๐	$100 \times 200 = 50$ ----- ๕๐๐	๑๐๐
ภาคกลาง	$100 \times 200 = 50$ ----- ๕๐๐	$100 \times 200 = 50$ ----- ๕๐๐	๑๐๐
ภาคใต้	$100 \times 200 = 50$ ----- ๕๐๐	$100 \times 200 = 50$ ----- ๕๐๐	๑๐๐
รวม	๒๐๐	๒๐๐	๔๐๐

๒) คำนวณหาค่าไคสแควร์

$$\begin{aligned}
 X^2 &= \frac{(40 - 50)^2}{50} + \frac{(60 - 50)^2}{50} + \frac{(30 - 50)^2}{50} + \\
 &\quad \frac{(70 - 50)^2}{50} + \frac{(50 - 50)^2}{50} + \frac{(50 - 50)^2}{50} + \\
 &\quad \frac{(80 - 50)^2}{50} + \frac{(20 - 50)^2}{50} \\
 &= 2 + 2 + 8 + 8 + 0 + 0 + 18 + 18 \\
 &= 56
 \end{aligned}$$

๕) พิจารณาผลการทดสอบสมมติฐาน

ผลการทดสอบสมมติฐาน พบว่า ค่า X^2 คำนวณ = ๕๖ ซึ่งมากกว่า X^2 ตาราง = ๗.๘๑๕ ซึ่งเป็นไปตามสมมติฐานที่ตั้งไว้ ดังนั้น จึงสรุปได้ว่า ภูมิภาค มีความสัมพันธ์กับความคิดเห็นต่อการปฏิรูปการเมือง หมายความว่า ประชาชนในแต่ละภูมิภาคมีความเห็นต่อการปฏิรูปการเมือง แตกต่างกัน จึงเป็นไปตามสมมติฐานที่ตั้งไว้

๕) สถิติวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's r)

ค่าสหสัมพันธ์ (Correlation) เป็นค่าสถิติที่ใช้บอกหรืออธิบายความสัมพันธ์ระหว่างตัวแปร 2 ตัว กล่าวคือ ถ้าตัวแปรตัวใดตัวหนึ่งมีค่าเปลี่ยนไป ตัวแปรอีกตัวหนึ่งจะมีค่าเปลี่ยนไปด้วย เรียกว่า ตัวแปรทั้งสองตัวนั้นมีความสัมพันธ์กัน ถ้าการเปลี่ยนแปลงนั้นเป็นไปในทิศทางเดียวกัน คือ เพิ่มขึ้นหรือลดลงตามกัน เรียกว่ามีความสัมพันธ์ทางบวก ถ้าตัวแปรเปลี่ยนในทิศทางที่สวนกัน คือ ตัวหนึ่งเพิ่มตัวหนึ่งลด เรียกว่าสัมพันธ์กันทางลบและสหสัมพันธ์เชิงศูนย์หรือความสัมพันธ์เท่ากับ 0 คือ ไม่มีความสัมพันธ์กัน ค่าความสัมพันธ์ที่คำนวณออกมาเป็นตัวเลข เรียกว่า ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) ซึ่งค่าสัมประสิทธิ์สหสัมพันธ์มีค่าตั้งแต่ ๑.๐๐ ถึง -๑.๐๐

ในการพิจารณาความสัมพันธ์พิจารณาได้จากหลักกว้างๆ ดังนี้

๑. ถ้าค่า r มีค่า ± ๐.๙๐ ถึง ๑.๐๐ หมายถึง มีความสัมพันธ์อย่างมาก
๒. ถ้าค่า r มีค่า ± ๐.๗๐ ถึง ๐.๙๐ หมายถึง มีความสัมพันธ์ค่อนข้างสูง
๓. ถ้าค่า r มีค่า ± ๐.๕๐ ถึง ๐.๗๐ หมายถึง มีความสัมพันธ์ปานกลาง
๔. ถ้าค่า r มีค่า ± ๐.๓๐ ถึง ๐.๕๐ หมายถึง มีความสัมพันธ์ค่อนข้างน้อย
๕. ถ้าค่า r มีค่า ± ๐.๐๐ ถึง ๐.๓๐ หมายถึง มีความสัมพันธ์น้อยมาก แทบไม่มีเลย

ดังนั้น เมื่อคำนวณหาความสัมพันธ์แล้ว ผู้วิจัยจะทราบความสัมพันธ์กันใน

๓ ส่วน คือ

- ๑) ทราบว่า ตัวแปรต้นและตัวแปรตาม มีความสัมพันธ์กันหรือไม่
- ๒) ทราบว่า มีความสัมพันธ์กันไปในทิศทางใด คือ ทางบวก หรือ ทางลบ
- ๓) ทราบว่า มีความสัมพันธ์กันอยู่ในระดับใด

การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson Product Moment Correlation (r_{xy})) เป็นการหาความสัมพันธ์ระหว่างตัวแปร ๒ ตัว หรือระหว่างข้อมูล ๒ ชุด โดยทั้ง ๒ ชุดจะต้องเป็นข้อมูลที่อยู่ในมาตราอันดับหรือมาตราส่วน เช่น เงินรายได้ อายุ เวลา น้ำหนัก ส่วนสูง เป็นต้น

ข้อตกลงเบื้องต้น

๑. ตัวแปรหรือข้อมูลทั้ง ๒ ชุด อยู่ในมาตราอันดับหรือมาตราส่วน
๒. ข้อมูลทั้ง ๒ ชุด มีการแจกแจงแบบปกติ และมีความสัมพันธ์เชิงเส้นตรง
๓. ข้อมูลในแต่ละชุดจะต้องมีความเป็นอิสระต่อกัน

ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน คำนวณจากสูตร

$$r_{xy} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{[n \sum X^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}}$$

เมื่อ r_{xy} เป็น ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน

$\sum X$ เป็น ผลรวมของข้อมูลที่วัดได้จากตัวแปรตัวที่ ๑ (X)

$\sum Y$ เป็น ผลรวมของข้อมูลที่วัดได้จากตัวแปรตัวที่ ๒ (Y)

$\sum XY$ เป็น ผลรวมของผลคูณระหว่างข้อมูลตัวแปรที่ ๑ และ ๒

$\sum X^2$ เป็น ผลรวมของกำลังสองของข้อมูลที่วัดได้จากตัวแปรตัวที่ ๑

$\sum Y^2$ เป็น ผลรวมของกำลังสองของข้อมูลที่วัดได้จากตัวแปรตัวที่ ๒

n เป็น ขนาดของกลุ่มตัวอย่าง

^{๑๓} ศุภชัย ยาวะประภาช, หน่วยที่ ๑๒ การวิเคราะห์ข้อมูลของรัฐประศาสตร์ ด้วยสถิติเชิงอนุมาน (๒), เอกสารชุดวิชา ขอบข่ายและวิธีใช้ทางรัฐประศาสนศาสตร์ หน่วยที่ ๙-๑๕, นนทบุรี: มหาวิทยาลัยสุโขทัยธรรมาธิราช, ๒๕๓๙, หน้า ๖๐๗.

ตัวอย่างเช่น

ในการศึกษาความสัมพันธ์ระหว่างประสบการณ์การทำงาน (ปี) กับระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน) ของพนักงานในหน่วยงานของรัฐแห่งหนึ่ง ผลปรากฏมีดังนี้

คนที่	ประสบการณ์การทำงาน (ปี)	ระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน)
๑	๑	๕
๒	๒	๓
๓	๓	๔
๔	๔	๔
๕	๕	๓
๖	๖	๓
๗	๗	๓
๘	๑๐	๓
รวม	๓๘	๒๘

จากตารางดังกล่าว จงทดสอบว่า ประสบการณ์การทำงาน (ปี) ของพนักงาน มีความสัมพันธ์กับระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน) ที่ระดับความเชื่อมั่น ๙๕% หรือไม่? มีความสัมพันธ์อยู่ในทิศทางใด? อยู่มีความสัมพันธ์อยู่ในระดับใด?

๑) ตั้งสมมติฐาน

$H_0 : \rho = 0$: ประสบการณ์การทำงาน (ปี) ไม่มีความสัมพันธ์กับระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน) (ตัวแปร X และ Y ไม่มีความสัมพันธ์กัน)

$H_1 : \rho \neq 0$: ประสบการณ์การทำงาน (ปี) มีความสัมพันธ์กับระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน) (ตัวแปร X และ Y มีความสัมพันธ์กัน)

๒) กำหนดระดับนัยสำคัญ (Level of Significance)

กำหนดนัยสำคัญทางสถิติที่ ๐.๐๕

๓) หาค่าวิกฤต (Critical Value : CV)

ผู้วิจัยหาค่าสถิติที่ใช้เป็นจุดแบ่งระหว่างบริเวณการยอมรับและบริเวณการปฏิเสธสมมติฐาน โดยการเปิดตารางค่าวิกฤตได้ โดยดูจากค่าชั้นของความอิสระ (Degree of Freedom) เมื่อเปิดตารางค่าวิกฤตของสัมประสิทธิ์สหสัมพันธ์ (r) แบบสองทาง ใน

ภาคผนวก หน้าที่ ๒๙๕ ที่ $df = n - ๒ = ๘ - ๒ = ๖$ ที่ช่องนัยสำคัญทางสถิติที่ ๐.๐๕ ได้ค่าเท่ากับ ๐.๘๑๑ ดังนั้น ค่า $r_{ตาราง} = ๐.๘๑๑$ หมายความว่า ค่า $r_{คำนวณ}$ มากกว่า $r_{ตาราง}$ จึงยอมรับสมมติฐาน กล่าวคือ ประสบการณ์การทำงาน (ปี) มีความสัมพันธ์กับระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน) แต่ถ้าค่า $r_{คำนวณ}$ น้อยกว่า $r_{ตาราง}$ จึงปฏิเสธสมมติฐาน กล่าวคือ ประสบการณ์การทำงาน (ปี) ไม่มีความสัมพันธ์กับระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน) หรือ

ผู้วิจัยอาจทดสอบหาค่านัยสำคัญโดยใช้สถิติทดสอบที่ ก็ได้ ดังนี้

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

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

๔) คำนวณโดยใช้สถิติ

คนที่	ประสบการณ์ การทำงาน (ปี) X	ระยะเวลาใน การสรุปรงาน ต่อชิ้นงาน (วัน) Y	X^2	Y^2	XY
๑	๑	๕	๑	๒๕	๕
๒	๒	๓	๔	๙	๖
๓	๓	๔	๙	๑๖	๑๒
๔	๔	๔	๑๖	๑๖	๑๖
๕	๕	๓	๒๕	๙	๑๕
๖	๖	๓	๓๖	๙	๑๘
๗	๗	๓	๔๙	๙	๒๑
๘	๑๐	๓	๑๐๐	๙	๓๐
รวม(Σ)	๓๘	๒๘	๒๔๐	๑๐๒	๑๒๓

จากสูตร

$$r_{xy} = \frac{n\sum XY - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}}$$

ดังนั้น

$$\begin{aligned}
 r_{xy} &= \frac{8(123) - (38)(28)}{\sqrt{[8(240) - 38^2][8(102) - 28^2]}} \\
 &= \frac{984 - 1064}{\sqrt{[1920 - 1444][816 - 784]}} \\
 &= \frac{-80}{\dots} \\
 &= 0.23.418 \\
 &= -0.648
 \end{aligned}$$

๕) พิจารณาผลการทดสอบสมมติฐาน

ผลการทดสอบสมมติฐาน พบว่า ค่า $r_{คำนวณ} = -0.648$ ซึ่งน้อยกว่า $r_{ตาราง} = -0.811$ จึงปฏิเสธสมมติฐาน กล่าวคือ ประสิทธิภาพการทำงาน (ปี) ไม่มีความสัมพันธ์กับระยะเวลาในการปฏิบัติงานต่อชิ้นงาน (วัน) ซึ่งไม่เป็นไปตามสมมติฐานที่ตั้งไว้ หรือ

ผู้วิจัยอาจทดสอบหาค่านัยสำคัญโดยใช้สถิติทดสอบที ก็ได้ ดังนี้

สถิติทดสอบ เป็นการทดสอบแบบสองทาง

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

โดยกำหนดระดับนัยสำคัญ .05

$$df = n-2 = 8 - 2 = 6$$

จุดตัดของค่าวิกฤตที่ 0.05 ในภาคผนวกหน้า ๒๘๕ คือ ค่าของ $t_{ตาราง}$ $df = 6$ ที่ ๒.๔๔๗ หมายความว่า หากค่า $t_{คำนวณ}$ มากกว่า $t_{ตาราง}$ จึงยอมรับสมมติฐานที่ตั้งไว้ แต่ถ้าหากค่า $t_{คำนวณ}$ น้อยกว่า $t_{ตาราง}$ จึงปฏิเสธสมมติฐานที่ตั้งไว้

คำนวณโดยใช้สถิติ

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

แทนค่า

$$\begin{aligned}
 t &= \frac{-0.648 \pm \sqrt{0.42 - 4(-0.648)(-0.648)}}{2(-0.648)} \\
 &= \frac{-0.648 \pm 0.361}{-1.296} \\
 &= -2.084
 \end{aligned}$$

พิจารณาผลการทดสอบสมมติฐาน

ผลการทดสอบสมมติฐาน พบว่า ค่า $t_{\text{คำนวณ}} = -2.084$ ซึ่งน้อยกว่า $t_{\text{ตาราง}} = 2.447$ ซึ่งไม่เป็นไปตามสมมติฐานที่ตั้งไว้ ดังนั้น จึงสรุปได้ว่า ประสิทธิภาพการทำงาน (ปี) ไม่มีความสัมพันธ์กับระยะเวลาในการสรุปรงานต่อชิ้นงาน (วัน)

๑๒.๖ การวิเคราะห์ข้อมูลด้วยคอมพิวเตอร์โปรแกรมสำเร็จรูป

การวิเคราะห์ข้อมูลทางสถิตินั้น เป็นการประมวลผลข้อมูลที่มีปริมาณมาก ดังนั้น การนำคอมพิวเตอร์มาช่วยในการวิเคราะห์ข้อมูล จะช่วยประหยัดเวลาและค่าใช้จ่าย มีความสะดวกมากกว่าที่จะคำนวณด้วยเครื่องคิดเลขทุกๆ ไป และในการใช้คอมพิวเตอร์ ส่วนใหญ่นักวิจัยนิยมที่จะใช้โปรแกรมสำเร็จรูป (Package Program) มากกว่าที่จะเขียนโปรแกรมขึ้นมาเอง โปรแกรมสำเร็จรูปที่นำมาใช้วิเคราะห์ข้อมูลจะเป็นโปรแกรมสำเร็จรูปทางสถิติโดยตรง แต่ในปัจจุบันมีโปรแกรมสำเร็จรูปที่ทำงานด้านอื่นๆ สามารถนำมาวิเคราะห์ข้อมูลทางสถิติได้ เช่น โปรแกรม EXCEL ซึ่งเป็นโปรแกรมประเภทกระดานขุด (Spread Sheet) แต่เนื่องจากโปรแกรมประเภทนี้ถูกสร้างขึ้นมาสำหรับงานบางอย่างที่ไม่ใช่สถิติโดยตรง ดังนั้น การวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติบางอย่าง โปรแกรมประเภทนี้อาจทำไม่ได้ หรือทำได้แต่อาจจะให้รายละเอียดน้อยกว่า ดังนั้นผู้วิจัยส่วนใหญ่จึงนิยมใช้โปรแกรมสำเร็จรูปทางสถิติมากกว่า แต่อย่างไรก็ตาม โปรแกรม EXCEL จะมีความสามารถในการนำเสนอข้อมูลด้วยกราฟหรือแผนภูมิ

การวิเคราะห์ข้อมูลสถิติ สามารถทำได้หลายแบบดังนี้

๑. การวิเคราะห์ด้วยมือ (Manual Data Analysis) เป็นการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติโดยการคำนวณเอง หรืออาจจะใช้เครื่องมือบางอย่างช่วยได้ เช่น เครื่องคิดเลข เป็นต้น วิธีนี้เหมาะสำหรับข้อมูลที่มีจำนวนไม่มากนัก การคำนวณไม่ยุ่งยาก

๒. การวิเคราะห์ด้วยเครื่องคอมพิวเตอร์ (Computer Data Analysis) เป็นการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติ โดยใช้เครื่องคอมพิวเตอร์มาช่วย ซึ่งผู้ใช้จำเป็นต้องใช้เครื่องคอมพิวเตอร์ได้และเลือกโปรแกรมวิเคราะห์ได้เหมาะสม วิธีนี้เหมาะสำหรับข้อมูลจำนวนมาก เป็นวิธีทางที่สะดวก รวดเร็ว แต่ต้องมีความระมัดระวังในเรื่องของข้อมูล กล่าวคือ ข้อมูลที่ใช้ในการวิเคราะห์ต้องผ่านการตรวจสอบความถูกต้องมาแล้ว

๑) ประเภทของโปรแกรม

การวิเคราะห์ข้อมูลสถิติที่มีจำนวนมากควรเลือกคอมพิวเตอร์ช่วยในการวิเคราะห์ ซึ่งต้องอาศัยโปรแกรม หรือ ซอฟต์แวร์มาช่วย

โปรแกรม หรือ ซอฟต์แวร์ หมายถึง ชุดของคำสั่งที่มีการจัดเรียงลำดับได้อย่างถูกต้อง ซึ่งสามารถทำงานและได้ผลลัพธ์ตามที่ผู้ใช้โปรแกรมต้องการ

โปรแกรมจำแนกได้เป็น ๒ ประเภท คือ

๑. โปรแกรมที่ผู้ใช้เขียนขึ้นเอง (User's Written Program) เป็นโปรแกรมที่ผู้ใช้เขียนสั่งให้คอมพิวเตอร์ทำงานได้ตามความต้องการ หรือ ตรงตามวัตถุประสงค์ และเหมาะสมด้วยภาษาคอมพิวเตอร์ โดยใช้เทคนิคและความชำนาญของผู้เขียนโปรแกรมด้วยภาษาคอมพิวเตอร์ ที่นิยมใช้ เช่น ภาษาเบสิก ภาษาซี เป็นต้น

๒. โปรแกรมสำเร็จรูป (Package Program) เป็นโปรแกรมที่มีผู้เขียนได้เขียนไว้เรียบร้อยแล้ว โปรแกรมสำเร็จรูปจะให้ความสะดวกในการใช้งานมาก โดยที่ผู้ใช้ไม่จำเป็นต้องมีความรู้ทางด้านคอมพิวเตอร์มากนัก เพียงแต่เรียนรู้วิธีการใช้งาน ซึ่งส่วนมากจะมีคำอธิบายการใช้โปรแกรมมาให้ และในขณะที่ทำงานก็สามารถขอรายละเอียดเพิ่มเติมได้ตลอดเวลาในการใช้โปรแกรมสำเร็จรูป

๒) โปรแกรมสำเร็จรูปทางสถิติ

โปรแกรมสำเร็จรูปทางสถิติ (Statistical Package) เป็นโปรแกรมสำหรับการวิเคราะห์ข้อมูลทางสถิติโดยตรง ในระยะแรกโปรแกรมสำเร็จรูปทางสถิติที่ใช้กับคอมพิวเตอร์มีไม่มากนักและมีใช้เฉพาะเครื่องคอมพิวเตอร์เมนเฟรม โปรแกรมที่มีชื่อเสียงและเป็นที่รู้จักกันดีคือ โปรแกรมสำเร็จรูป SPSS (Statistical Package for the Social Sciences) โปรแกรมสำเร็จรูป SAS (Statistical Analysis System) และโปรแกรมสำเร็จรูป BMDP (Biomedical Computer Program)

ในปัจจุบันนี้ได้มีความก้าวหน้าทางเทคโนโลยีด้านไมโครคอมพิวเตอร์ ทำให้การประมวลผลข้อมูลมีประสิทธิภาพมากขึ้น ดังนั้นนักวิจัยจึงหันมาใช้ไมโครคอมพิวเตอร์มากขึ้น เพราะมีความสะดวกมากกว่า จึงทำให้มีผู้ผลิตโปรแกรมสำเร็จรูปทางสถิติ สำหรับเครื่องไมโครคอมพิวเตอร์มากขึ้น ซึ่งมีทั้งโปรแกรมขนาดเล็กที่วิเคราะห์สถิติเฉพาะอย่าง และโปรแกรมขนาดใหญ่ที่สามารถวิเคราะห์สถิติหลายประเภท ได้แก่ โปรแกรม SPSS SAS และ BMDP ได้ถูกพัฒนามาใช้กับเครื่องไมโครคอมพิวเตอร์เช่นเดียวกัน ซึ่งมีชื่อต่างๆ กันดังนี้

โปรแกรม SPSS สำหรับเครื่องไมโครคอมพิวเตอร์มีชื่อเรียกว่า SPSS/PC +, SPSS FOR WINDOWS

โปรแกรม SAS สำหรับเครื่องไมโครคอมพิวเตอร์มีชื่อเรียกว่า SAS ON PC DOS

โปรแกรม BMDP สำหรับเครื่องไมโครคอมพิวเตอร์มีชื่อเรียกว่า BMDP-PC

โปรแกรมสำเร็จรูปทางสถิติอาจจำแนกตามองค์ประกอบของโปรแกรมได้เป็น ๒ ประเภทใหญ่ๆ คือ

๑. โปรแกรมเบ็ดเสร็จ เป็นโปรแกรมที่เขียนขึ้นสำหรับวิเคราะห์ข้อมูลทางสถิติแบบอเนกประสงค์ มีความสามารถในการวิเคราะห์ข้อมูลเชิงสถิติทุกประเภท โดยทั่วไปแล้วจะเป็นโปรแกรมที่มีขนาดใหญ่ เช่น โปรแกรม SPSS/PC, SAS, BMDP-PC, GENSTAT, MINITAB และ STATA เป็นต้น

๒. โปรแกรมที่เน้นเฉพาะด้าน เป็นโปรแกรมที่เขียนขึ้นเพื่อเน้นการวิเคราะห์เชิงสถิติประเภทใดประเภทหนึ่งโดยเฉพาะ เช่น

- โปรแกรมสำหรับกำหนดการเชิงเส้น (Linear Programming) เช่น โปรแกรม LINDO โปรแกรม GINO โปรแกรม LINGO

- โปรแกรมที่เน้นการพยากรณ์ของข้อมูลอนุกรมเวลา เช่น โปรแกรม TSP (Time Series Program)

- โปรแกรมที่เน้นทางด้าน Statistical Modeling เช่น โปรแกรม GLIM

ลักษณะการใช้งานของโปรแกรมทางสถิติ อาจแบ่งออกได้เป็น ๒ ประเภทใหญ่ๆ คือ

๑. โปรแกรมชนิดสั่งให้ทำงานด้วยคำสั่ง (Command Driven) ก่อนที่จะใช้โปรแกรมชนิดนี้ผู้ใช้จะต้องศึกษาคำสั่ง (Commands) ต่างๆ ให้เข้าใจเสียก่อน จากนั้นจึงเขียนชุดคำสั่งเพื่อทำการวิเคราะห์ข้อมูล ตัวอย่างของโปรแกรมประเภทนี้ เช่น SPSS/PC , MINITAB , SAS และ STATA เป็นต้น

๒. โปรแกรมชนิดสั่งให้ทำงานโดยอาศัยเมนู (Menu Driven) โปรแกรมประเภทนี้จะใช้งานง่ายกว่าโปรแกรมชนิดคำสั่ง จึงเหมาะสำหรับผู้เริ่มใช้โปรแกรมสำเร็จรูปทางสถิติ ทั้งนี้เนื่องจาก โปรแกรมประเภทนี้จะแสดงรายการต่างๆ ออกทางหน้าจอ แล้วให้ผู้ใช้เลือกว่าต้องการใช้รายการใด เช่น โปรแกรม STATPACK , SPSS for Windows^{๑๓}

๑๒.๗ สรุปท้ายบท

การวิเคราะห์ข้อมูล หมายถึง การนำข้อมูลที่เก็บรวบรวมได้มาจำแนกหรือจัดให้เป็นหมวดหมู่ จัดระเบียบ คำนวณค่าสถิติต่างๆ สรุปผลและนำเสนอให้อยู่ในรูปแบบที่สั้นกะทัดรัด โดยแบ่งออกเป็น ๒ ประเภท คือ ๑) การวิเคราะห์ข้อมูลเชิงคุณภาพ แบ่งเป็น ๒ ลักษณะ ได้แก่ ๑) การวิเคราะห์ข้อมูลภาคสนามหรือจากปรากฏการณ์ และ ๒) การวิเคราะห์ข้อมูลจากเอกสาร หรือ การวิเคราะห์เนื้อหา ๒) การวิเคราะห์ข้อมูลเชิงปริมาณ อันเป็นการนำข้อมูลที่ให้ค่าเป็นจำนวนตัวเลขไปวิเคราะห์

กระบวนการการวิเคราะห์ข้อมูล มี ๕ กระบวนการ คือ การเก็บรวบรวมข้อมูล, การแยกประเภทข้อมูลและจัดชั้น, การสังเขป, การหาข้อสรุป และการพยากรณ์

วิธีการวิเคราะห์ข้อมูล ประกอบด้วย

๑.การวิเคราะห์ข้อมูลเชิงคุณภาพ เริ่มตั้งแต่ ๑.๑) การตรวจสอบข้อมูลก่อนทำการวิเคราะห์ประกอบด้วย ๑) การตรวจสอบข้อมูลสามเส้า (Data Triangulation) ด้านเวลา สถานที่ และบุคคล ๒) การตรวจสอบสามเส้าด้านผู้วิจัย (Investigator Triangulation) ๓) การตรวจสอบสามเส้าด้านทฤษฎี (Theory Triangulation) และ ๔) การตรวจสอบสามเส้าด้านวิธีรวบรวมข้อมูล (Methodological Triangulation) ๑.๒ วิธีการวิเคราะห์ข้อมูล แบ่งออกเป็น ๒ ส่วน คือ ๑) การวิเคราะห์ข้อมูลแบบสร้างข้อสรุป มีอยู่ด้วยกัน ๓ ชนิด คือ ๑.การวิเคราะห์แบบอุปนัย ๒. การจำแนกชนิดข้อมูล มีทั้งวิธีแบบใช้ทฤษฎี และแบบไม่ใช้ทฤษฎี ๓. การเปรียบเทียบข้อมูล และ ๒) การวิเคราะห์เนื้อหา (Content Analysis) มีองค์ประกอบหลัก ๖ ประการ ได้แก่ ๑. แหล่งที่มาของข้อความ หรือสาระ ได้แก่ ผู้สื่อ (Source/Sender) ๒. กระบวนการใส่ความของสาระ (Encoding Process) ๓. ตัวสาระและข้อความ (Message) ๔. วิธีถ่ายทอดสาระไปยังผู้อื่น (Channel of Transmission) ๕. ผู้รับสาระ (Detector) ๖. กระบวนการถอดความหมายของสาระ (Decoding Process) ขั้นตอนในการวิเคราะห์เนื้อหา คือ ๑) ผู้วิจัยจะต้องตั้งกฎเกณฑ์ขึ้น ๒) ผู้วิจัยจะต้องวางเค้าโครงของ

^{๑๓} http://tsl.tsu.ac.th/courseware/math2/lesson1/less1_4.htm

ข้อมูล ๓) ผู้วิจัยจะต้องคำนึงถึง บริบท (Context) หรือสภาพแวดล้อม ๔) การวิเคราะห์เนื้อหาจะทำตามเนื้อหาที่ปรากฏ (Manifest Content) ๕) การวัดความถี่ของการใช้ หรือ การใช้การวิเคราะห์เชิงปริมาณ แล้วให้ได้คำตอบที่มีความหมายสัมพันธ์กับสิ่งที่ต้องการ ค้นหา

๒. การวิเคราะห์ข้อมูลเชิงปริมาณ แบ่งออกเป็น ๒ ประเภท คือ ๑) การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา หมายถึง การใช้กระบวนการทางสถิติที่ใช้บรรยายลักษณะของข้อมูลที่ผู้วิจัยเก็บรวบรวมจากประชากรหรือกลุ่มตัวอย่างที่สนใจทั้งที่เป็นข้อมูลเชิงคุณภาพ และปริมาณ โดยผลที่ได้จากการศึกษาจะไม่สามารถนำไปอ้างอิงไปยังกลุ่มตัวอย่างหรือกลุ่มประชากรอื่นได้ การนำเสนอข้อมูลดังกล่าว อาจเป็นการนำเสนอด้วยตาราง ร้อยละ รูปกราฟประเภทต่างๆ และสถิติที่ใช้ ได้แก่ การแจกแจงความถี่ การวัดแนวโน้มสู่ส่วนกลาง (ค่าเฉลี่ยเลขคณิต, มัธยฐาน, ฐานนิยม) การวัดการกระจาย (ส่วนเบี่ยงเบนมาตรฐาน, ความแปรปรวน) เป็นต้น ๒) การวิเคราะห์ข้อมูลโดยใช้สถิติเชิงอ้างอิง หมายถึง การใช้กระบวนการทางสถิติที่ใช้อธิบายคุณลักษณะของสิ่งที่ต้องการวิจัยในกลุ่มใดกลุ่มหนึ่ง และสามารถอ้างอิงไปถึงประชากรที่มีค่าจริงหรือไปยังกลุ่มอื่นๆ ได้ โดยกลุ่มที่นำมาศึกษาจะต้องเป็นตัวแทนที่ดีของประชากร จะเรียกว่า กลุ่มตัวอย่าง และสถิติที่ใช้ ได้แก่ สถิติค่าที่ สถิติวิเคราะห์ความแปรปรวนทางเดียว สถิติวิเคราะห์ไคสแควร์ และสถิติวิเคราะห์ความสัมพันธ์ของเพียร์สัน เป็นต้น

๒.๑ วิธีการวิเคราะห์ข้อมูลโดยใช้สถิติเชิงพรรณนา ประกอบด้วย ๑) การแจกแจงความถี่ มี ๒ ประเภท คือ ๑)การแจกแจงของข้อมูลแบบไม่จัดกลุ่ม ๒) การแจกแจงของข้อมูลแบบจัดเป็นกลุ่ม ๒) การวัดแนวโน้มเข้าสู่ส่วนกลาง มี ๓ วิธี คือ ค่าเฉลี่ยเลขคณิต มัธยฐาน ฐานนิยม ๓) การวัดการกระจาย สถิติที่นิยมใช้ ได้แก่ พิสัย ค่าเบี่ยงเบนมาตรฐาน และค่าความแปรปรวน

๒.๒ วิธีการวิเคราะห์ข้อมูลโดยใช้สถิติเชิงอ้างอิง (Inference Statistics) ซึ่งมีขั้นตอนการวิเคราะห์ข้อมูลโดยใช้สถิติอ้างอิง ประกอบด้วย ๑) สังเกตเพื่อระบุปัญหา ๒) ขั้่นตั้งสมมติฐาน ๓) ขั้่นการรวบรวมข้อมูล ๔)ขั้่นสรุปผล และสถิติที่ใช้ในการวิเคราะห์ข้อมูลเชิงอ้างอิงประกอบด้วย ๑) สถิติวิเคราะห์ค่าที่ (t-test) ๒) สถิติวิเคราะห์ความแปรปรวนทางเดียว (One Way ANOVA) ๓) สถิติวิเคราะห์ไค-สแควร์ (Chi-Square Test) ๔) สถิติวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน (Pearson's r)

การวิเคราะห์ข้อมูลด้วยคอมพิวเตอร์โปรแกรมสำเร็จรูป เป็นการประมวลผลข้อมูลที่มีปริมาณมาก จึงนำคอมพิวเตอร์มาช่วยในการวิเคราะห์ข้อมูล จะช่วยประหยัดเวลาและค่าใช้จ่าย มีความสะดวกมากกว่าที่จะคำนวณด้วยเครื่องคิดเลขทั่วไป และในการใช้คอมพิวเตอร์ ส่วนใหญ่นักวิจัยนิยมที่จะใช้โปรแกรมสำเร็จรูป (Package Program) มากกว่าที่จะเขียนโปรแกรมขึ้นมาเอง โปรแกรมสำเร็จรูปที่นำมาใช้วิเคราะห์ข้อมูลจะเป็นโปรแกรมสำเร็จรูปทางสถิติโดยตรง

การวิเคราะห์ข้อมูลสถิติ สามารถทำได้หลายแบบดังนี้

๑. การวิเคราะห์ด้วยมือ (Manual Data Analysis)
๒. การวิเคราะห์ด้วยเครื่องคอมพิวเตอร์ (Computer Data Analysis)

คำถามท้ายบท

๑. จงอธิบายกระบวนการวิเคราะห์ข้อมูลเชิงคุณภาพมาพอเข้าใจ?
๒. คะแนนของนิสิตในรายวิชาระเบียบวิธีวิจัยทางรัฐศาสตร์ จำนวน ๑๐ คน ดังนี้ ๕๕ ๕๙ ๖๑ ๖๕ ๖๖ ๗๔ ๘๒ ๘๒ ๙๕ ๙๖ จงหาค่าเฉลี่ย มัธยฐาน ฐานนิยม
๓. พนักงานในองค์การบริหารส่วนตำบล จำนวน ๘ คน มีประสบการณ์ในการทำงานเกี่ยวกับท้องถิ่น ดังนี้ ๓ ปี, ๖ ปี, ๗ ปี, ๔ ปี, ๒ ปี, ๖ ปี, ๕ ปี ๔ ปี จงหาส่วนเบี่ยงเบนมาตรฐาน
๔. พนักงานองค์การบริหารส่วนตำบล ๕๐ คน มีความคิดเห็นเกี่ยวกับการปฏิรูปการเมือง โดยเฉลี่ยเท่ากับ ๓.๕ ส่วนเบี่ยงเบนมาตรฐานอยู่ที่ ๑.๓๔ เมื่อจำแนกตามเพศ มีค่าเฉลี่ยดังนี้

เพศ	จำนวน	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน
ชาย	๓๐	๔.๐๐	๑.๓๐๐
หญิง	๒๐	๓.๐๐	๑.๕๖๕
รวม	๕๐	๓.๕๐	๑.๓๔๐

จากข้อมูลดังกล่าว เปรียบเทียบว่า พนักงานมีความคิดเห็นต่อการปฏิรูปการเมืองแตกต่างกันหรือไม่

๕. การศึกษาความคิดเห็นของประชาชนต่อการปฏิรูปการเมือง โดยทำการศึกษาจากประชาชนในเขตเทศบาลของประเทศไทย จำนวน ๒๐๐ คน โดยจำแนกเป็นเขตพื้นที่ได้ดังนี้

สภาพพื้นที่	ความคิดเห็นของประชาชนต่อการปฏิรูปการเมือง		รวม
	เห็นด้วย	ไม่เห็นด้วย	
ในเขตเทศบาล	๖๐	๔๐	๑๐๐
นอกเขตเทศบาล	๓๐	๗๐	๑๐๐
รวม	๙๐	๑๑๐	๒๐๐

จากตารางดังกล่าว จงทดสอบว่าประชาชนในแต่ละพื้นที่มีความเห็นต่อการปฏิรูปการเมืองแตกต่างกันหรือไม่ ที่ระดับความเชื่อมั่น ๙๕%