บทที่ 5
การวิเคราะห์ในสถิติอิงพารามิเตอร์

1. แนวคิดเกี่ยวกับสถิติอิงพารามิเตอร์

สถิติเชิงอนุมานหรือสถิติสรุปอ้างอิง (inferential statistics) ใช้สำหรับอนุมานผลลัพธ์ที่ได้จากกลุ่มตัวอย่างซึ่งเป็นตัวแทนที่ดีของประชากรหนึ่ง อ้างอิงกลับไปสู่ประชากรนั้นเอง

วิธีการวิเคราะห์ในสถิติอนุมานเหมาะสมกับข้อมูลของกลุ่มตัวอย่างที่ได้มาโดยวิธีการสุ่ม หรืออาศัยหลักการของความน่าจะเป็นในการคัดเลือกตัวอย่างเพื่อให้เป็นตัวแทนที่ดีของประชากรหนึ่ง และอัตราการตอบข้อมูลของกลุ่มตัวอย่างต้องสูงมาก ๆ คือ มีข้อมูลที่สูญหายไปหรือไม่ตอบน้อยมาก ๆ ดังนั้นวิธีการวิเคราะห์ในสถิติอนุมานจึงไม่เหมาะสมกับข้อมูลของกลุ่มตัวอย่างที่ไม่ได้มาโดยวิธีการสุ่ม และข้อมูลได้มาจากประชากรทั้งหมดอยู่แล้ว

ในการวิจัยทางสังคมศาสตร์ส่วนใหญ่เครื่องมือที่ใช้ในการเก็บรวบรวมข้อมูลมักเป็นแบบสอบถามที่สร้างขึ้นตามทฤษฎี การเก็บรวบรวมข้อมูลโดยการแจกแบบสอบถามไปยังกลุ่มตัวอย่างแล้วเก็บคืนมักจะได้ข้อมูลตอบกลับไม่ครบทั้งหมด ซึ่งเป็นปัญหาที่ทราบกันดี ดังนั้นถ้าข้อมูลส่วนที่ไม่ตอบกลับมานี้มีจำนวนมาก หรือเป็นบางส่วนของกลุ่มตัวอย่างที่มีลักษณะเดียวกัน เช่น ช่วงอายุ , เพศ , สถานะทางเศรษฐกิจ อาจทำให้คุณลักษณะของกลุ่มตัวอย่างที่เหลือบิดเบือนไปไม่ถูกต้องตามคุณลักษณะของประชากรนั้น ตัวอย่างเช่น ผู้ที่ไม่ตอบแบบสอบถามส่วนใหญ่เป็นคนสูงอายุ ก็จะทำให้การแจกแจงอายุของกลุ่มตัวอย่างถูกบิดเบือนไปและมีผลต่อความถูกต้องของการประมาณอายุเฉลี่ยของประชากร

1.1 การทดสอบแบบอิงพารามิเตอร์และแบบไม่อิงพารามิเตอร์ (Parametric and Non-Parametric Tests)

ลักษณะของตัวแปรที่สนใจศึกษา มี 2 แบบใหญ่ ๆ คือ ตัวแปรแบบจำแนกประเภท (categorical variable) ซึ่งมีสเกลการวัดแบบแบ่งประเภทและแบบอันดับ และตัวแปร

แบบต่อเนื่อง ซึ่งมีสเกลการวัดแบบช่วง และแบบอัตราส่วน ลักษณะของตัวแปรจะเป็นตัวกำหนดในการเลือกใช้วิธีการวิเคราะห์ทางสถิติที่เหมาะสม นอกจากนี้ยังสามารถแบ่งวิธีการวิเคราะห์ทางสถิติออกเป็น 2 ประเภท คือ สถิติอิงพารามิเตอร์ (parametric statistics) และสถิติไม่อิงพารามิเตอร์ (non-parametric statistics) โดยเน้นที่คุณลักษณะของประชากรมากกว่าสเกลการวัดของข้อมูล

การทดสอบนัยสำคัญบางชนิดใช้ได้เฉพาะเมื่อการแจกแจงของตัวแปรในประชากรที่มีการแจกแจงโดยประมาณเป็นแบบปกติเท่านั้น บางการทดสอบใช้ในการเปรียบเทียบกลุ่มตัวอย่างซึ่งมาจากประชากรที่แตกต่างกันนั้นคือ ต้องมีการกระจายเท่ากัน หรือมีความแปรปรวนเท่ากัน เราเรียกการทดสอบประเภทนี้ว่า การทดสอบแบบอิงพารามิเตอร์ และถ้าเราไม่สามารถบอกได้ว่ากลุ่มตัวอย่างสุ่มมาจากประชากรซึ่งมีคุณลักษณะตามข้อตกลงเบื้องต้นที่กำหนดเกี่ยวกับการแจกแจงของประชากรตามที่ต้องการของวิธีการทดสอบทางสถิติแบบอิงพารามิเตอร์ เราก็ต้องใช้วิธีการทดสอบทางสถิติอีกประเภทหนึ่งคือ การทดสอบแบบไม่อิงพารามิเตอร์ หรืออาจเรียกว่าการทดสอบที่เป็นอิสระจากการแจกแจงของประชากร (Distribution-free Tests) อย่างไรก็ตามการเลือกใช้วิธีการทางสถิติในการทดสอบทั้ง 2 ประเภทนี้ ก็ยังขึ้นอยู่กับชนิดของตัวแปรด้วย ตัวอย่างเช่น เมื่อตัวแปรที่สนใจศึกษาเป็นตัวแปรแบบจำแนกประเภท (categorical) เราไม่สามารถบอกเกี่ยวกับลักษณะการแจกแจงของประชากรข้ามระหว่างประเภทต่าง ๆ ได้ เราจึงจำเป็นต้องเลือกใช้วิธีการวิเคราะห์ที่เป็นอิสระจากการแจกแจงของประชากร

นอกจากนี้ในกรณีที่ประชากรมีการแจกแจงไม่เป็นแบบปกติตามข้อตกลงเบื้องต้นก็ยังมีวิธีการแปลง (transform) ข้อมูล เพื่อให้มีการแจกแจงแบบปกติ หรือการสร้างตัวแปรดัมมี (dummy variables) จากข้อมูลจำแนกประเภทแล้วอนุญาตให้ใช้วิธีการแบบอิงพารามิเตอร์ ตัวอย่างเช่น การวิเคราะห์การถดถอย (regression analysis)

ในการตัดสินใจว่าคุณลักษณะของประชากรเป็นไปตามข้อตกลงเบื้องต้นหรือไม่นั้น ถ้าไม่มีข้อสรุปลักษณะของประชากรตามที่ต้องการ เราก็ต้องทำการตรวจสอบจากคุณลักษณะของกลุ่มตัวอย่างแทน โดยมีความเชื่อมั่นว่ากลุ่มตัวอย่างนี้เป็นตัวแทนที่ดีของประชากร

1.2 การประมาณค่าพารามิเตอร์ (estimation of parameters)

ในการสรุปอ้างอิงเกี่ยวกับพารามิเตอร์ของประชากรจะเกี่ยวข้องกับความไม่แน่นอน เพราะสรุปอ้างอิงมาจากกลุ่มตัวอย่างกลุ่มหนึ่งเท่านั้น สถิติเชิงอนุมานจึงเกี่ยวข้องกับความไม่แน่นอน ซึ่งเป็นแนวคิดเกี่ยวกับความน่าจะเป็นและการแจกแจงของการสุ่มตัวอย่างของตัวสถิติ (sampling distribution of the statistic)

สถิติเชิงอนุมานมีเนื้อหาเกี่ยวกับ 2 ประเด็นใหญ่ ๆ คือ ประเด็นแรก การประมาณค่าพารามิเตอร์ของประชากร และประเด็นที่สอง การทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ของประชากร โดยใช้ข้อมูลตัวอย่างเป็นหลักฐานในการสนับสนุนหรือปฏิเสธสมมติฐานทางสถิติเกี่ยวกับค่าพารามิเตอร์ของประชากรซึ่งไม่ทราบค่า

การประมาณค่าพารามิเตอร์ของประชากรมี 2 แบบ คือ การประมาณค่าแบบจุด (point estimation) และการประมาณค่าแบบช่วง (interval estimation) การประมาณค่าแบบจุด จะได้ตัวสถิติเป็นค่าตัวเลขตัวหนึ่ง และส่วนเบี่ยงเบนมาตรฐานของตัวสถิติตัวนั้น ตัวอย่างเช่น การประมาณค่าเฉลี่ยของประชากร μ จากตัวอย่างเป็นตัวสถิติ คือ ค่าเฉลี่ยของตัวอย่าง $\bar{x}$ ส่วนการประมาณค่าแบบช่วงจะได้ช่วงของค่าสถิติที่บอกเป็นค่าขอบเขตต่ำ (lower bound) และค่าขอบเขตสูง (upper bound)

1.2.1 การประมาณค่าแบบจุด

สมมติว่าสุ่มตัวอย่างนักเรียน 55 คน ที่เรียนวิชาหลักการวางแผนการทดลอง ได้คะแนนสอบดังในตารางที่ 4.2 คำนวณค่าเฉลี่ยของคะแนนสอบของตัวอย่างได้ $\bar{x} = 45.6$ คะแนน และส่วนเบี่ยงเบนมาตรฐาน $s = 9.24$ คะแนน เราเรียก $\bar{x}$ ว่าค่าประมาณแบบจุด ใช้ในการประมาณค่าพารามิเตอร์ μ ซึ่งค่าส่วนเบี่ยงเบนมาตรฐานหรืออาจเรียกว่าความคลาดเคลื่อน มาตรฐานของตัวสถิติ $\bar{x}$ จะเป็นตัวบอกความถูกต้องของการประมาณค่าตัวสถิตินั้น

1.2.2 การประมาณค่าแบบช่วง

เนื่องจากเราไม่ทราบค่าจริงของพารามิเตอร์ เราต้องการจะหาช่วงซึ่งคลุมค่าจริงของพารามิเตอร์ โดยอาศัยหลักความน่าจะเป็น ความน่าจะเป็นนี้เรียกว่าระดับความเชื่อมั่น (level of confidence) หมายถึง ระดับความน่าจะเป็นที่กำหนดขึ้น เช่น 95 เปอร์เซ็นต์ เป็นค่าที่แสดงความเชื่อมั่น เกี่ยวกับการประมาณค่าพารามิเตอร์ของประชากรว่ามี

ความถูกต้อง 95 เปอร์เซ็นต์ มีโอกาสผิดเพียง 5 เปอร์เซ็นต์ ส่วนช่วงของค่าที่คลุมค่าสถิติที่คำนวณได้จากกลุ่มตัวอย่าง ซึ่งคาดว่าค่าพารามิเตอร์ของประชากรจะตกอยู่ในช่วงนี้ หมายถึง ช่วงความเชื่อมั่น (confidence interval) ทั้งระดับความเชื่อมั่นและช่วงความเชื่อมั่นนี้ขึ้นอยู่กับขนาดของกลุ่มตัวอย่างและวิธีการสุ่มตัวอย่าง ถ้าใช้กลุ่มตัวอย่างขนาดใหญ่มากขึ้นเพียงใดก็จะได้ระดับความเชื่อมั่นสูงขึ้น และช่วงความเชื่อมั่นแคบมากขึ้นเพียงนั้น

การคำนวณหาช่วงความเชื่อมั่นของค่าเฉลี่ยของประชากร μ ต้องประมาณค่าความคลาดเคลื่อนมาตรฐาน (standard error) ของ $\bar{x}$ ก่อน มักแทนด้วย se มีสูตรการคำนวณคือ

$$se = \frac{s}{\sqrt{n}}$$

เมื่อ s คือ ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

n คือ ขนาดของกลุ่มตัวอย่าง

การคำนวณช่วงความเชื่อมั่น หรือแทนด้วย CI เราต้องมีการกำหนดระดับความเชื่อมั่นก่อน มีสูตรการคำนวณคือ

$$CI = \bar{x} \pm (z \times se)$$

เมื่อ $z = 1.96$ ถ้ากำหนดให้ระดับความเชื่อมั่นเท่ากับ 95 เปอร์เซ็นต์ และ

$z = 2.58$ ถ้ากำหนดให้ระดับความเชื่อมั่นเท่ากับ 99 เปอร์เซ็นต์ ซึ่งได้จากตารางการ

แจกแจงแบบปกติ

ตัวอย่างเช่น สุ่มตัวอย่างนักเรียน 55 คน จากประชากรที่มีการแจกแจงแบบปกติ และไม่ทราบค่าเบี่ยงเบนมาตรฐาน σ ได้ข้อมูลคะแนนสอบดังตารางที่ 4.2 คำนวณหาค่าเฉลี่ยของตัวอย่าง $\bar{x} = 45.6$ คะแนน และส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง $s = 9.24$ คะแนน

ต้องการคำนวณหา 95% ช่วงความเชื่อมั่นของคะแนนเฉลี่ยของประชากร คือ

$$95\%\ \text{CI ของ}\ \mu = \left(45.6 - 1.96\frac{9.24}{\sqrt{55}}\ ,\ 45.6 + 1.96\frac{9.24}{\sqrt{55}}\right)$$

$$= (43.16\ ,\ 48.04)\ \text{คะแนน}$$

การแปลความหมายของช่วงความเชื่อมั่น ถ้าสุ่มตัวอย่างขนาดเท่า ๆ กันมา 100 ครั้งอย่างเป็นอิสระกัน จากประชากรที่มีการแจกแจงแบบปกติมาหาช่วงความเชื่อมั่นของพารามิเตอร์ทุกครั้ง 95% ของช่วงความเชื่อมั่นนี้จะคลุมค่าจริงของพารามิเตอร์ที่ไม่ทราบค่า นั่นคือเชื่อว่าค่าพารามิเตอร์ของประชากรจะอยู่ในช่วงความเชื่อมั่นนี้ ด้วยความน่าจะเป็นเท่ากับ 95% มีโอกาสเพียง 5% เท่านั้นที่ค่าพารามิเตอร์ของประชากรจะไม่อยู่ในช่วงความเชื่อมั่นนี้

1.3 การทดสอบสมมติฐานทางสถิติ 9testing statistical hypotheses)

ในการวิจัยทางสังคมศาสตร์มีสมมติฐาน 2 ชนิด คือ สมมติฐานทางทฤษฎี (theoretical hypotheses) และสมมติฐานทางสถิติ (statistical hypotheses) สมมติฐานทางทฤษฎีคือ แนวโน้มของคำตอบของปัญหาที่ทำการวิจัยนั้นว่าควรจะมีคำตอบอย่างไร โดยวิเคราะห์ตามทฤษฎีเท่านั้นไม่ต้องใช้หลักการทางสถิติมาช่วยอธิบาย

สมมติฐานทางสถิติมี 2 รูปแบบคือ สมมติฐานสูญ (null hypotheses) มักแทนด้วย H_0 และสมมติฐานแย้ง (alternative hypotheses) มักแทนด้วย H_1 สมมติฐานทางสถิติถูกใช้ในกระบวนการของการอนุมานผลลัพธ์ที่ได้จากกลุ่มตัวอย่างเพื่อไปสู่ประชากรทั่วไป การทดสอบสมมติฐานเป็นการใช้ข้อมูลของกลุ่มตัวอย่างเพื่อเป็นหลักฐานในการตัดสินใจระหว่างสมมติฐานสูญและสมมติฐานแย้ง ในขั้นตอนของการทดสอบสมมติฐานมีการพิจารณาที่สำคัญ 2 อย่างคือ (1) พิจารณาเกี่ยวกับจำนวนชั้นอิสระ (degree of freedom) ของสถิติทดสอบ (test statistics) และ (2) ทำการเปรียบเทียบค่าสถิติทดสอบที่คำนวณได้กับค่าวิกฤตที่เปิดจากตาราง

จำนวนชั้นอิสระ หมายถึง จำนวนของค่าของข้อมูลที่เป็นอิสระในการผันแปรได้เมื่อค่าของข้อมูลทั้งหมดถูกจำกัดด้วยเงื่อนไขที่วางไว้ให้ ตัวอย่างเช่น ข้อมูลในตารางขนาด 2×2 จำนวนชั้นอิสระจะเกี่ยวข้องกับจำนวนของช่องในตารางคือ เรารู้ขนาดของตัวอย่างและผลรวม

ส่วนริมของตาราง ดังนั้นถ้าเราทราบข้อมูลที่เป็นจำนวนนับของช่องหนึ่งในตาราง เราจะสามารถคำนวณหาข้อมูลที่เป็นจำนวนนับในช่องที่เหลือของตารางได้ นั่นคือมีข้อมูลเพียงช่องเดียวเท่านั้นในตารางที่เป็นอิสระสามารถผันแปรค่าได้ และถ้ากำหนดให้ช่องนี้มีค่าที่แน่นอนค่าหนึ่ง จะทำให้อีก 3 ช่องที่เหลือถูกกำหนดค่าตามช่องแรกนั้น จะได้ว่าในตารางขนาด 2×2 จะมีจำนวนชั้นอิสระเท่ากับ 1

การดำเนินการทดสอบสมมติฐานมีขั้นตอนพื้นฐาน 5 ขั้นตอน คือ

1. กำหนดสมมติฐานสูญ และสมมติฐานแย้ง ตามการทดสอบที่ต้องการคือ อาจเป็นแบบทดสอบทางเดียว (one-tailed hypotheses or one-sided hypotheses) หรือแบบสองทาง (two-tailed hypotheses or two-sided hypotheses)

2. กำหนดระดับนัยสำคัญ (level of significance)

3. เลือกสถิติทดสอบที่เหมาะสมกับคุณลักษณะของข้อมูล ซึ่งได้แก่ สเกลการวัดข้อมูล และเทคนิควิธีการวิเคราะห์ทางสถิติ ตัวอย่างเช่น สถิติ Z สถิติ t สถิติ F หรือสถิติไคสแคว์

4. คำนวณค่าของสถิติทดสอบ

5. ทำการตัดสินใจว่าจะปฏิเสธหรือไม่ปฏิเสธสมมติฐานสูญ

1.3.1 การกำหนดสมมติฐานสูญ และสมมติฐานแย้ง

การกำหนดสมมติฐานทางเดียว หมายถึง กำหนดให้ค่าของพารามิเตอร์ในสมมติฐานแย้งตกอยู่ในทางด้านเดียวของค่าของพารามิเตอร์ที่อยู่ในสมมติฐานสูญ และการทดสอบสมมติฐานทางเดียวเรียกว่า การทดสอบทางเดียว (one-sided tests หรือ one-tailed tests)

ตัวอย่างเช่น ผู้สอนต้องการประเมินความสามารถของนักเรียนที่เรียนวิชาหลักการวางแผนการทดลอง ให้พารามิเตอร์ μ เป็นคะแนนเฉลี่ยของประชากร ผู้สอนคาดว่า μ ต่ำกว่า 50 คะแนน ทำให้กำหนดสมมติฐานแย้ง $H_1 : \mu < 50$ ถ้าผู้สอนคาดผิดในสถานการณ์นี้อาจเป็น $\mu = 50$ และ $\mu > 50$ ด้วยเหตุผลนี้เราอาจกำหนดสมมติฐานสูญได้อย่างง่าย ๆ คือ $\mu = 50$ นั่นคือกำหนดสมมติฐานสูญและสมมติฐานแย้งได้คือ

$$H_0 : \mu = 50 \quad \text{คู่กับ} \quad H_1 : \mu < 50$$

การกำหนดสมมติฐานสองทาง หมายถึง กำหนดให้ค่าของพารามิเตอร์ในสมมติฐานแย้งตกอยู่ทั้ง 2 ด้านของค่าพารามิเตอร์ที่อยู่ในสมมติฐานสูญ และการทดสอบสมมติฐานสองทาง เรียกว่า การทดสอบสองทาง (two-sided tests หรือ two-tailed tests)

ตัวอย่างเช่น ผู้สอนต้องการประเมินความสามารถของนักเรียนที่เรียนวิชาหลักการวางแผนการทดลอง โดยสุ่มตัวอย่างนักเรียน 55 คน ได้ข้อมูลคะแนนสอบของนักเรียนเหล่านี้ดังในตารางที่ 4.2 ผู้สอนอยากทราบว่าจากข้อมูลคะแนนสอบของกลุ่มตัวอย่างบอกได้หรือไม่ว่าคะแนนเฉลี่ยของประชากรแตกต่างจาก 50 คะแนน นั่นคือผู้สอนกำลังหาหลักฐานที่จะสนับสนุนว่า $\mu \neq 50$ ดังนั้นจึงกำหนดสมมติฐานสูญ และสมมติฐานแย้งได้คือ

$$H_0 : \mu = 50 \quad \text{คู่กับ} \quad H_1 : \mu \neq 50$$

คะแนนเฉลี่ยของประชากรแตกต่างจาก 50 คะแนน นั่นคือผู้สอนกำลังหาหลักฐานที่จะสนับสนุนว่า $\mu \neq 50$ ดังนั้นจึงกำหนดสมมติฐานสูญ และสมมติฐานแย้งได้คือ

$$H_0 : \mu = 50 \quad \text{คู่กับ} \quad H_1 : \mu \neq 50$$

1.3.2 กำหนดระดับนัยสำคัญ

ระดับนัยสำคัญ (level of significance) คือโอกาสหรือความน่าจะเป็นของการตัดสินใจผิดพลาดในการวิเคราะห์ข้อมูลเมื่อทำการทดสอบสมมติฐานทางสถิติ นั่นคือ เมื่อทำการอนุมานเป็นข้อสรุปเกี่ยวกับพารามิเตอร์ของประชากรจากผลลัพธ์ที่ได้ของกลุ่มตัวอย่างซึ่งเป็นตัวแทนของประชากรนั้น เรามักแทนระดับนัยสำคัญด้วย α ซึ่งสามารถเขียนในรูปของความน่าจะเป็นได้คือ

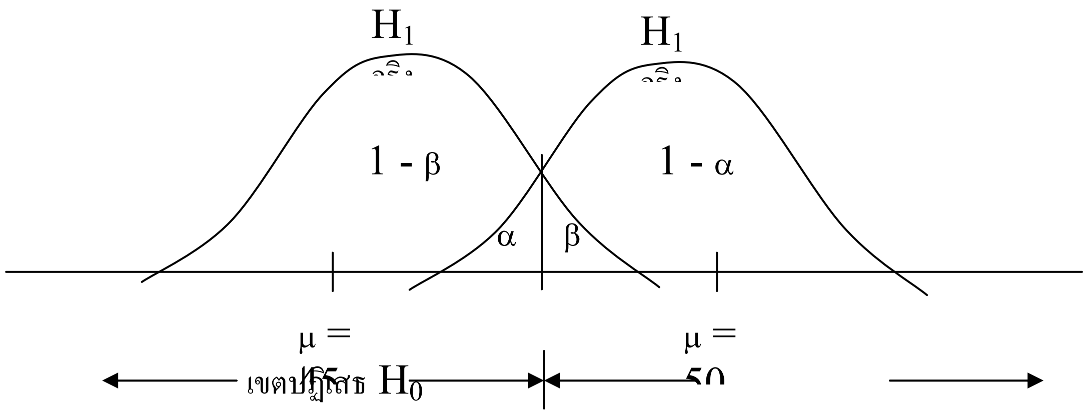
$$\begin{aligned} \alpha &= \text{ความน่าจะเป็นของการตัดสินใจผิดพลาดประเภท I} \\ &= P\ (\text{Type I error}) \end{aligned}$$

การตัดสินใจผิดพลาดประเภท I คือ การตัดสินใจปฏิเสธ H_0 เมื่อ H_0 เป็นจริง และการตัดสินใจผิดพลาดประเภท II คือ การตัดสินใจไม่ปฏิเสธ H_0 เมื่อ H_0 ไม่เป็นจริง เรามักแทนด้วย β สามารถเขียนในรูปของความน่าจะเป็นได้คือ

$$\beta = P\text{ (Type II error)}$$

จากทฤษฎีความน่าจะเป็นเรารู้ว่ามีความเป็นไปได้ที่กลุ่มตัวอย่างอาจไม่เป็นตัวแทนที่ดีของประชากร และมีความเป็นไปได้ที่อาจทำให้ผลการวิเคราะห์ค่าสถิติอยู่ที่หางของโค้งการแจกแจงแบบปกติของตัวอย่าง ดังนั้นเราจึงต้องระวังในการตัดสินใจเกี่ยวกับสมมติฐานทางสถิติ โดยทำให้มีความสมดุลย์ระหว่างความคลาดเคลื่อนประเภท I และความคลาดเคลื่อนประเภท II โดยการกำหนดระดับความเชื่อมั่น $(1 - \alpha)$ โดยทั่วไปเรามักจะกำหนดระดับความเชื่อมั่นที่ระดับ $(1 - \alpha) \times 100$ เท่ากับ 95% หรือ 99% โดยมีเหตุผลที่สำคัญคือต้องการหลีกเลี่ยงความผิดพลาดประเภท I โดยกำหนดให้ $\alpha = .05$ หรือ $.01$

ภาพที่ 5.1 แสดงความน่าจะเป็นของความคลาดเคลื่อนประเภท I ภายใต้พื้นที่ใต้โค้งของการแจกแจงแบบปกติที่มีค่าพารามิเตอร์ $\mu = 50$ และความน่าจะเป็นของความคลาดเคลื่อนประเภท II ภายใต้พื้นที่ใต้โค้งของการแจกแจงแบบปกติที่มีค่าพารามิเตอร์ $\mu = 45$ ในกรณีที่สมมติฐานแย้งเป็นจริง



ภาพที่ 5.1

1.3.2 กำหนดระดับนัยสำคัญ

ระดับนัยสำคัญ (level of significance) คือโอกาสหรือความน่าจะเป็นของการตัดสินใจผิดพลาดในการวิเคราะห์ข้อมูลเมื่อทำการทดสอบสมมติฐานทางสถิติ นั่นคือ เมื่อทำการอนุมานเป็นข้อสรุปเกี่ยวกับพารามิเตอร์ของประชากรจากผลลัพธ์ที่ได้ของกลุ่มตัวอย่าง ซึ่งเป็นตัวแทนของประชากรนั้น เรามักแทนระดับนัยสำคัญด้วย α ซึ่งสามารถเขียนในรูปของความน่าจะเป็นได้คือ

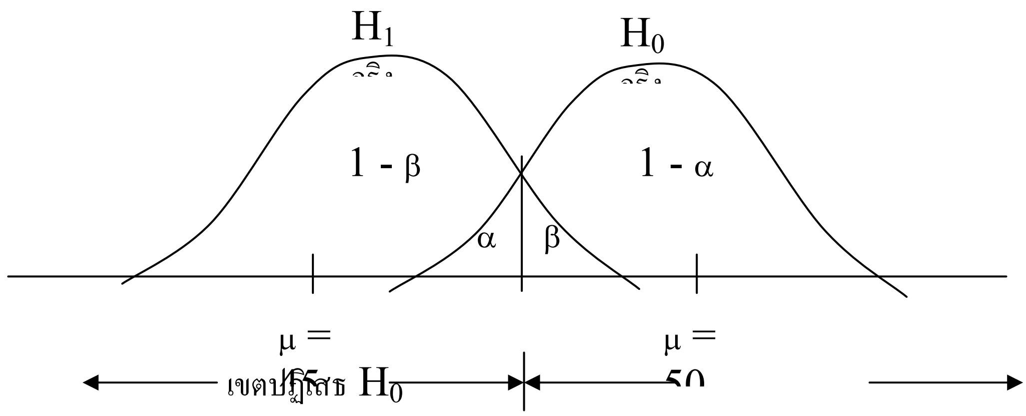
$$\begin{aligned}\alpha &= \text{ความน่าจะเป็นของการตัดสินใจผิดพลาดประเภท I} \\ &= P\text{ (Type I error)}\end{aligned}$$

การตัดสินใจผิดพลาดประเภท I คือ การตัดสินใจปฏิเสธ H_0 เมื่อ H_0 เป็นจริง และการตัดสินใจผิดพลาดประเภท II คือ การตัดสินใจไม่ปฏิเสธ H_0 เมื่อ H_0 ไม่เป็นจริง เรามักแทนด้วย β สามารถเขียนในรูปของความน่าจะเป็นได้คือ

$$\beta = P\text{ (Type II error)}$$

จากทฤษฎีความน่าจะเป็นเรารู้ว่ามีความเป็นไปได้ที่กลุ่มตัวอย่างอาจไม่เป็นตัวแทนที่ดีของประชากร และมีความเป็นไปได้ที่อาจทำให้ผลการวิเคราะห์ค่าสถิติอยู่ที่หางของโค้งการแจกแจงแบบปกติของตัวอย่าง ดังนั้นเราจึงต้องระวังในการตัดสินใจเกี่ยวกับสมมติฐานทางสถิติ โดยทำให้มีความสมดุลย์ระหว่างความคลาดเคลื่อนประเภท I และความคลาดเคลื่อนประเภท II โดยการกำหนดระดับความเชื่อมั่น $(1 - \alpha)$ โดยทั่วไปเรามักจะกำหนดระดับความเชื่อมั่นที่ระดับ $(1 - \alpha) \times 100$ เท่ากับ 95% หรือ 99% โดยมีเหตุผลที่สำคัญคือต้องการหลีกเลี่ยงความผิดพลาดประเภท I โดยกำหนดให้ $\alpha = .05$ หรือ $.01$

ภาพที่ 5.1 แสดงความน่าจะเป็นของความคลาดเคลื่อนประเภท I ภายใต้พื้นที่ใต้โค้งของการแจกแจงแบบปกติที่มีค่าพารามิเตอร์ $\mu = 50$ และความน่าจะเป็นของความคลาดเคลื่อนประเภท II ภายใต้พื้นที่ใต้โค้งของการแจกแจงแบบปกติที่มีค่าพารามิเตอร์ $\mu = 45$ ในกรณีที่สมมติฐานแย้งเป็นจริง



ภาพที่ 5.1

1.3.3 เลือกสถิติทดสอบ

ในการเลือกสถิติทดสอบ ต้องพิจารณาจากคุณลักษณะของข้อมูลและข้อตกลงเบื้องต้นของสถิติทดสอบ ตัวอย่างเช่น ในสถานการณ์ของตัวอย่างข้างต้นผู้สอนวิชาหลักการวางแผนการทดลองต้องการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร 1 กลุ่ม เรามักเลือกใช้สถิติทดสอบ t-test ซึ่งจะอธิบายรายละเอียดเกี่ยวกับสถิติทดสอบ t-test ในตอนต่อไป ส่วนสถิติทดสอบแบบอื่น ๆ ก็จะอธิบายในบทอื่น ๆ

1.3.4 คำนวณค่าของสถิติทดสอบ

ถ้าคะแนนเฉลี่ยของกลุ่มตัวอย่าง $\bar{x}$ ที่คำนวณจากกลุ่มตัวอย่างนักเรียน 55 คน ที่มาจากประชากรซึ่งมีการแจกแจงแบบปกติและไม่ทราบค่า σ ซึ่งเราจะใช้เป็นฐานสำหรับการตัดสินใจปฏิเสธหรือไม่ปฏิเสธ H_0 ได้ว่า $\bar{x} = 45.6$ คะแนน และ $s = 9.24$ คะแนน สถิติทดสอบคือ

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$= \frac{45.6 - 50}{\frac{9.24}{\sqrt{55}}} = \frac{-4.4}{1.245}$$

$$= -3.534$$

1.3.5 การตัดสินใจว่าจะปฏิเสธหรือไม่ปฏิเสธ H_0

จากสมมติฐานทางเดียว $H_0 : \mu = 50$ คู่กับ $H_1 : \mu < 50$

กำหนดระดับนัยสำคัญ $\alpha = .05$

วิธีการทดสอบสมมติฐานทำได้โดยการเปรียบเทียบ ค่าที่คำนวณได้ t โดยไม่คำนึงถึงเครื่องหมายมักเขียนว่า $|t|$ กับค่า t_α ที่เปิดจากตารางที่ $df = n - 1$ ถ้า $|t| \leq t_\alpha$ จะตัดสินใจปฏิเสธ H_0 เนื่องจากค่าที่เปิดจากตาราง $t_{.05, 54} = 1.684$ ทำให้ค่าที่คำนวณได้ $|t| = 3.534$ มากกว่าค่าที่เปิดจากตาราง จึงตัดสินใจไม่ปฏิเสธ $H_0 : \mu = 50$

2. การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร 1 กลุ่ม

สถิติที่ใช้ในการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร 1 กลุ่ม คือ t – test ซึ่งมีข้อกำหนด ดังนี้

ข้อตกลงเบื้องต้น (Assumption) ของการทดสอบ t – test คือ

(1) ประชากรมีการแจกแจงแบบปกติ หรือขนาดตัวอย่างน้อยกว่าหรือเท่ากับ 30

(2) ตัวแปรมีสเกลการวัดแบบช่วง หรือแบบอัตราส่วน

2.1 การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากรกลุ่มเดียว

ข้อตกลงเบื้องต้น คือ กลุ่มตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติและไม่ทราบความแปรปรวนของประชากร

เรามักใช้ t – test ในการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ μ ของการแจกแจงแบบปกติ เรากำหนดสมมติฐานสูญว่า $H_0 : \mu = \mu_0$ คู่กับสมมติฐานแย้ง $H_1 : \mu \neq \mu_0$ การทดสอบจะเทียบกับค่าสถิติ t ซึ่งคือฟังก์ชันของค่าของกลุ่มตัวอย่างขนาด n ที่คำนวณจากสูตร คือ

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

ที่มีการแจกแจงแบบปกติมีค่าเฉลี่ยเท่ากับ 0 และความแปรปรวนเท่ากับ $df / (df - 2)$ เมื่อ $df = n - 1$ ถ้า $df = 30$ แล้ว $df / (df - 2) = 30/(30 - 2) = 1.07$ ซึ่งมีค่าเข้าใกล้ 1 และถ้า df มีค่ามาก ๆ $df / (df - 2)$ จะยิ่งมีค่าเข้าใกล้ 1 หมายความว่า ถ้า n มีขนาดใหญ่ขึ้น การแจกแจงของ t จะเข้าใกล้การแจกแจงแบบปกติ

วิธีการทดสอบทำได้ง่าย ๆ โดยการใช้ตารางเปรียบเทียบกับค่าที่คำนวณได้ t โดยไม่คำนึงถึงเครื่องหมาย มักจะเขียนว่า $|t|$ กับค่า t_α ที่ได้จากตารางเมื่อ H_0 เป็นจริง

$$\Pr(|t| \ge t_\alpha) = \alpha$$

เมื่อ α คือระดับนัยสำคัญ ในทางปฏิบัติเรามักจะเลือก α ที่ระดับ 0.05 , 0.01 หรือ 0.001 ถ้าได้ว่า $|t| \ge t_\alpha$ จะสรุปผลว่ามีนัยสำคัญที่ความน่าจะเป็น α ถ้าผลการคำนวณพบว่ามีนัยสำคัญที่กำหนด เราจะสรุปว่าปฏิเสธสมมติฐาน H_0 ที่ระดับนัยสำคัญนั้น แต่ถ้าผลการคำนวณไม่มีนัยสำคัญเราจะสรุปว่าไม่มีนัยสำคัญคือ ยอมรับ H_0

ปัจจุบันซอฟแวร์ทางสถิติสำหรับ t – test ที่ใช้ทดสอบสมมติฐานข้างต้นจะให้ผลลัพธ์เป็นความน่าจะเป็นจริง ๆ เมื่อ H_0 เป็นจริง ค่าของ $|t|$ เท่ากับหรือมากกว่าค่าที่สังเกตได้ ความน่าจะเป็นนี้เรียกว่า P-value คือ การวัดความเข้มข้นของหลักฐานที่จะคัดค้าน H_0 โดยอาศัยหลักฐานจากข้อมูล ถ้าค่า P-value ยิ่งเล็ก หมายถึง ยิ่งมีหลักฐานมากขึ้นในการคัดค้าน H_0

การทดสอบสมมติฐาน $H_0 : \mu = \mu_0$ คู่กับ $H_1 : \mu \neq \mu_0$ เรียกว่าการทดสอบสองทาง เพราะว่ามีเขตวิกฤตอยู่ทั้งสองด้านของค่าสถิติ t ค่าที่อยู่ปลายทางด้านบวกของ t แสดงนัยว่า $\mu > \mu_0$ และค่าที่อยู่ปลายทางด้านลบของ t แสดงนัยว่า $\mu < \mu_0$

เราสามารถทำการทดสอบแบบทางเดียวโดยที่เขตวิกฤติจะอยู่ทางปลายด้านใดด้านหนึ่งเท่านั้น เช่น สมมติฐานที่ต้องการทดสอบคือ $H_0 : \mu = \mu_0$ คู่กับ $H_1 : \mu > \mu_0$ หรือ $H_0 : \mu \le \mu_0$ คู่กับ $H_1 : \mu > \mu_0$ หรือ $H_0 : \mu \ge \mu_0$ คู่กับ $H_1 : \mu < \mu_0$

2.2 การใช้คำสั่ง One – Sample T Test ...

ตัวอย่างเช่น ผู้จัดการในบริษัทแห่งหนึ่งอ้างว่าเงินเดือนเฉลี่ยของพนักงานในบริษัทสูงกว่า 6,000 บาท ผู้วิจัยจึงทำการเก็บตัวอย่างจากพนักงานในบริษัทจำนวน 20 คน ดังข้อมูลในแฟ้มข้อมูล DataTest5.sav ที่มีตัวแปร salary ซึ่งมีสเกลการวัดแบบอัตราส่วน เพื่อทำการทดสอบว่าคำกล่าวอ้างของผู้จัดการเป็นจริงหรือไม่ มีขั้นตอนการใช้คำสั่ง ดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Compare Means, One Sample T Test... จะได้หน้าต่าง One – Sample T Test

2. ในหน้าต่าง One - Sample T Test ในช่องซ้ายมือ คลิกที่ตัวแปร salary แล้วคลิกที่หัวลูกศร ➢หน้าช่อง Test Variable(S) : ตัวแปร salary จะย้ายเข้าไปอยู่ในช่อง Test Variable(S) :

ในช่อง Test Value : ใส่ตัวเลข 6,000

คลิกที่ปุ่ม Options... จะได้หน้าต่าง One – Sample T Test : Options ดังภาพที่ 5.1

3. ในหน้าต่าง One – Sample T Test : Options

คำสั่ง Confidence Interval : เป็นคำสั่งสำหรับระบุช่วงความเชื่อมั่นของการทดสอบสมมติฐานทางสถิติ โดยปกติโปรแกรมจะกำหนดให้แล้วเท่ากับ 95% ถ้าต้องการเปลี่ยนก็สามารถใส่ค่าที่ต้องการได้ เช่น 99%

คำสั่ง Missing Values เป็นคำสั่งจัดการกับค่าที่ขาดหายไป (Missing) คือ

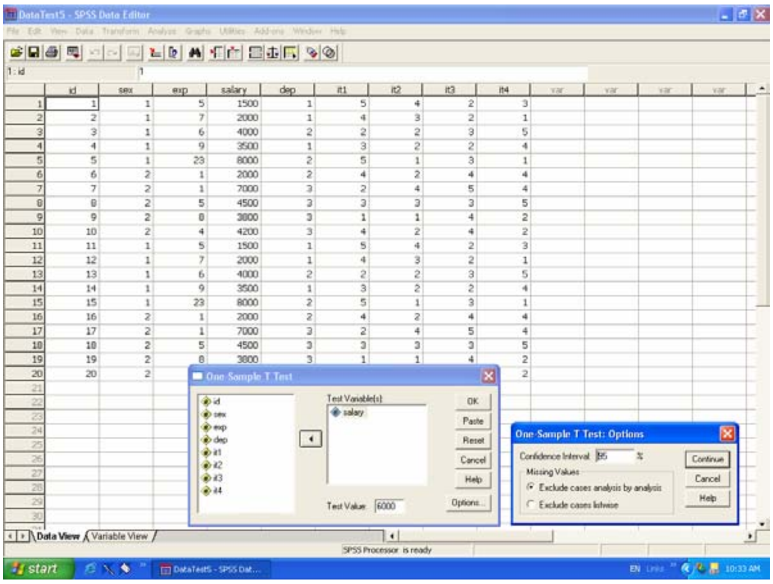
- ถ้าเลือกคำสั่ง O Exclude cases analysis by analysis ผลลัพธ์ที่ได้จะไม่นำตัวอย่างที่มีค่า missing ของตัวแปร salary มารวมในการวิเคราะห์ โดยปกติโปรแกรมจะเลือกคำสั่งนี้อยู่แล้ว

- ถ้าเลือกคำสั่ง O Exclude cases listwise ผลลัพธ์ที่ได้จะไม่นำตัวอย่างที่มีค่า missing ของตัวแปร salary มารวมในการวิเคราะห์เฉพาะกรณีที่มีการวิเคราะห์ครั้งละหลายตัวแปรพร้อมกันเท่านั้น โดยจำนวนตัวอย่างของทุกตัวแปรจะมีจำนวนเท่ากัน แต่ถ้าเป็นการวิเคราะห์ครั้งละ 1 ตัวแปร คำสั่งนี้จะให้ผลลัพธ์เหมือนคำสั่งแรก

แล้วคลิกปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

4. ในหน้าต่าง One – Sample T Test

คลิกปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 5.2



ภาพที่ 5.1

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
salary	20	4050.00	2046.949	457.712

One-Sample Test

	Test Value = 6000					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper
salary	-4.260	19	.000	-1950.000	-2908.00	-992.00

ภาพที่ 5.2

จากภาพผลลัพธ์มีตัวอย่างที่นำมาวิเคราะห์จำนวนทั้งหมด 20 คน กลุ่มตัวอย่างนี้มีเงินเดือนเฉลี่ย 4,050.00 บาท ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2,046.949

ตัวอย่างนี้เป็นการทดสอบแบบทางเดียว (One – tailed Testing) ซึ่งมีเขตวิกฤตอยู่ทางด้านขวา แต่ค่า t มีค่าเป็นลบซึ่งอยู่ด้านซ้าย ดังนั้น P-value จึงเริ่มต้นจากหางด้านขวาของพื้นที่ใต้โค้งไปถึง ณ ตำแหน่งของค่า t ซึ่งอยู่ด้านซ้าย จึงได้ค่า P-value เท่ากับ 1-sig (2 – tailed)/2 = 1 – (.002/2) ซึ่งมากกว่าระดับนัยสำคัญที่กำหนด $\alpha = .05$ จึงยอมรับ $H_0 : \mu \leq 6000$ หมายความว่า เงินเดือนเฉลี่ยของพนักงานน้อยกว่าหรือเท่ากับ 6,000 บาท

3. การทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร 2 กลุ่ม

มีวัตถุประสงค์เพื่อเปรียบเทียบค่าเฉลี่ยของประชากร 2 กลุ่มที่เป็นอิสระกันว่าแตกต่างกันหรือไม่

ข้อตกลงเบื้องต้นคือ

กลุ่มตัวอย่างทั้ง 2 กลุ่ม สุ่มมาจากประชากรที่มีการแจกแจงแบบปกติและไม่ทราบความแปรปรวนของประชากร

การทดสอบสมมติฐาน

สำหรับการทดสอบแบบสองทาง สมมติฐานที่ต้องการทดสอบคือ $H_0 : \mu_1 = \mu_2$ คู่กับ $H_1 : \mu_1 \neq \mu_2$ สำหรับการทดสอบแบบทางเดียว สมมติฐานที่ต้องการทดสอบคือ $H_0 : \mu \leq \mu_2$ คู่กับ $H_1 : \mu_1 > \mu_2$ หรือ $H_0 : \mu_1 \geq \mu_2$ คู่กับ $H_1 : \mu_1 < \mu_2$

ในการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร 2 กลุ่ม ควรต้องพิจารณาเกี่ยวกับความเป็นอิสระของประชากรทั้ง 2 กลุ่มนั้น เพื่อเลือกใช้การทดสอบได้ถูกต้อง มี 2 กรณี คือ

(1) ประชากร 2 กลุ่ม เป็นอิสระกัน

(2) ประชากร 2 กลุ่ม ไม่เป็นอิสระต่อกัน

3.1 การทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่ม ที่เป็นอิสระกัน

การเลือกใช้สถิติทดสอบ t–test ต้องพิจารณาที่ความแปรปรวนของประชากรทั้ง 2 กลุ่ม คือ

(1) ถ้าความแปรปรวนของประชากรทั้ง 2 กลุ่มเท่ากัน $(\sigma_1^2 = \sigma_2^2)$ แล้วให้ใช้ความแปรปรวน (pooled variance) ดังนั้นสถิติทดสอบ คือ

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2\,(1/n_1 + 1/n_2)}}$$

ที่มี $df = n1 + n2 - 2$

เมื่อ $\bar{x}_1$ = ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 1

$\bar{x}_2$ = ค่าเฉลี่ยของกลุ่มตัวอย่างที่ 2

$s_p^2 = [(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2]/(n_1 + n_2 - 2)$

s_1^2 = ความแปรปรวนของกลุ่มตัวอย่างที่ 1

s_2^2 = ความแปรปรวนของกลุ่มตัวอย่างที่ 2

n_1 = ขนาดของกลุ่มตัวอย่างที่ 1

n_2 = ขนาดของกลุ่มตัวอย่างที่ 2

(2) ถ้าความแปรปรวนของประชากรทั้ง 2 กลุ่ม ไม่เท่ากัน $(\sigma_1^2 \neq \sigma_2^2)$ แล้วให้ใช้ความแปรปรวนแยก (separate variance) ดังนั้นสถิติทดสอบคือ

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

ที่มี
$$df = \frac{[(s_1^2/n^1) + (s_2^2/n^2)]^2}{[(s_1^2/n_1)^2/(n_1 - 1)] + [(s_2^2/n_2)^2/(n_2 - 1)]}$$

3.2 การใช้คำสั่ง Independent – Samples T Test....

ตัวอย่างเช่น ต้องการทดสอบว่าพนักงานชายมีเงินเดือนเฉลี่ยน้อยกว่าพนักงานหญิงหรือไม่ จากแฟ้มข้อมูล DataTeat5.sav มีขั้นตอนการใช้คำสั่ง ดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Compare Means, Independent – Samples T Test จะได้หน้าต่าง Independent – Samples T Test

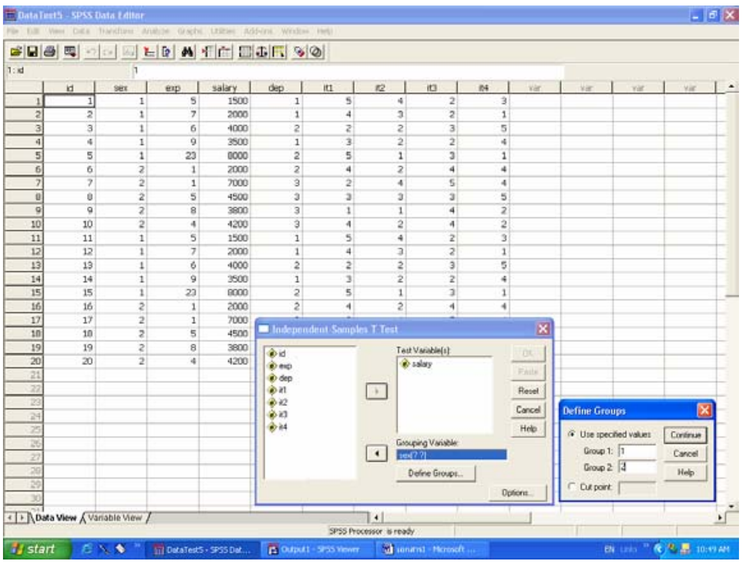
2. ในหน้าต่าง Independent – Samples T Test ในช่องซ้ายมือ คลิกที่ตัวแปร salary แล้วคลิกที่หัวลูกศร ➢ หน้าช่อง Test Variable(S) : ตัวแปร salary จะย้ายเข้าไปอยู่ในช่อง Test Variable(S) :

คลิกที่ตัวแปร sex แล้วคลิกที่หัวลูกศร ➢ หน้าช่อง Grouping Variable : ตัวแปร sex จะย้ายเข้าไปอยู่ในช่องนี้ แล้วคลิกที่ปุ่ม Define Groups... จะได้หน้าต่าง Define Groups ใส่เลข 1 ในช่อง Group 1: และใส่เลข 2 ในช่อง Group 2: ดังภาพที่ 5.3 แล้วคลิกที่ปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

ที่ปุ่ม Options... จะมีคำสั่งต่าง ๆ เหมือนกันกับการใช้คำสั่ง One – Sample T Test ในหัวข้อ 2.2 ทำเช่นเดียวกับที่อธิบายแล้ว

3. ในหน้าต่าง Independent – Samples T Test

คลิกปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 5.4



ภาพที่ 5.3

T-Test

Group Statistics

	sex	N	Mean	Std. Deviation	Std. Error Mean
salary	female	10	3800.00	2417.529	764.490
	male	10	4300.00	1691.810	534.997

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
salary	Equal variances assumed	.955	.341	-.536	18	.599	-500.000	933.095	-2460	1460.4
	Equal variances not assumed			-.536	16.110	.599	-500.000	933.095	-2477	1477.0

ภาพที่ 5.4

จากภาพผลลัพธ์มีตัวอย่างที่นำมาวิเคราะห์จำนวนทั้งหมด 20 คน เป็นพนักงานชาย 10 คน พนักงานหญิง 10 คน พนักงานชายมีเงินเดือนเฉลี่ยเท่ากับ 4,300.00 บาท ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1,691.810 ส่วนพนักงานหญิงมีเงินเดือนเฉลี่ย 3,800.00 บาท ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2,417.529

ผลการทดสอบ t – test สำหรับประชากร 2 กลุ่มที่เป็นอิสระกัน ทำการทดสอบสมมติฐานแบบทางเดียว ต้องพิจารณาที่ความแปรปรวนของประชากรทั้ง 2 กลุ่ม ก่อนว่าเท่ากันหรือไม่ โดยดูในคอลัมน์ Levene's test of Equality of Variance ดูที่ค่า Sig. เท่ากับ .341 ซึ่งมากกว่าระดับนัยสำคัญ $\alpha = .05$ ดังนั้นจึงยอมรับ H_0 : $\sigma_1^2 = \sigma_2^2$ นั่นคือสรุปว่าความแปรปรวนของเงินเดือนของประชากร 2 กลุ่มเท่ากัน

ดังนั้นจึงดูที่แถวแรกคือ Equal variances assumed แล้วดูค่าสถิติ t = -.536 ที่มี $df = n_1 + n_2 - 2 = 10 + 10 - 2 = 18$ ค่าสถิติ t มีค่าเป็นลบอยู่ทางด้านซ้าย ตัวอย่างนี้เป็นการทดสอบแบบทางเดียว สมมติฐานที่ต้องการทดสอบคือ H_0 : $\mu_1 \geq \mu_2$ คู่กับ H_1 : $\mu_1 < \mu_2$ ซึ่งมีเขตวิกฤติอยู่ทางด้านซ้าย ดังนั้น P-value จึงเริ่มต้นจากหางด้านซ้ายของพื้นที่ใต้โค้งไปถึง ณ ตำแหน่งของค่า t ซึ่งอยู่ด้านซ้าย จึงได้ค่า P-value เท่ากับ Sig. (2-tailed)/2 = .599/2 = .299 ซึ่งมากกว่าระดับนัยสำคัญที่กำหนด α = .05 จึงยอมรับ $H_0 : \mu_1 \geq \mu_2$ นั่นคือ พนักงานชายมีเงินเดือนเฉลี่ยไม่น้อยกว่าพนักงานหญิง

ในทางตรงข้าม ถ้าต้องการทดสอบสมมติฐานที่ว่าพนักงานชายมีเงินเดือนเฉลี่ยสูงกว่าพนักงานหญิงหรือไม่ สมมติฐานที่ต้องการทดสอบคือ $H_0 : \mu_1 \leq \mu_2$ คู่กับ $H_1 : \mu_1 > \mu_2$ ซึ่งมีเขตวิกฤติอยู่ทางด้านขวา ดังนั้น P-value จึงเริ่มต้นจากหางด้านขวาไปถึง ณ ตำแหน่งที่ค่า t อยู่คือด้านซ้าย จึงได้ค่า P-value เท่ากับ 1 – Sig. (2 – taited)/2 = 1 - .599/2 = .701 ซึ่งมากกว่า .05 จึงยอมรับ H0 นั่นคือพนักงานชายมีเงินเดือนเฉลี่ยไม่มากกว่าพนักงานหญิง ซึ่งดูเหมือนว่าผลสรุปจากการทดสอบสมมติฐานทั้ง 2 ครั้งข้างต้นขัดแย้งกัน แต่อาจเป็นไปได้ว่าพนักงานชายและหญิงมีเงินเดือนเฉลี่ยไม่แตกต่างกัน ซึ่งเป็นการทดสอบแบบสองทาง สมมติฐานที่ต้องการทดสอบคือ $H_0 : \mu_1 = \mu_2$ คู่กับ $H_1 : \mu_1 \neq \mu_2$ พิจารณาจากค่า Sig.(2 – tailed) = .599 มากกว่า .05 จึงยอมรับ $H_0 : \mu_1 = 2$ นั่นคือ พนักงานชายและหญิงมีเงินเดือนเฉลี่ยเท่ากันหรือไม่แตกต่างกัน หรือแตกต่างกันอย่างไม่มีนัยสำคัญ

3.3 การทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากร 2 กลุ่มที่ไม่เป็นอิสระกัน

กลุ่มตัวอย่างที่ไม่เป็นอิสระกัน (dependent samples) หมายถึง การเป็นสมาชิกของตัวอย่างกลุ่มหนึ่งขึ้นกับการเป็นสมาชิกของตัวอย่างอีกกลุ่มหนึ่ง เช่น การให้นักเรียนกลุ่มเดียวกันทำแบบทดสอบ 2 ฉบับ คือ แบบทดสอบก่อนเรียนและแบบทดสอบหลังเรียน เป็นต้น ส่วนตัวอย่างชนิดคู่ (paired samples) อาจได้มาจากคนฝาแฝด สัตว์ทดลองที่มาจากครอกเดียวกัน นักเรียนที่มี IQ เท่ากัน คนไข้ที่ป่วยเป็นโรคเดียวกันที่มีอาการเหมือนกัน เป็นต้น แล้วแบ่งตัวอย่างออกเป็น 2 กลุ่ม แบบจับคู่ ให้กลุ่มที่ 1 ได้รับทรีทเมนต์ 1 ให้กลุ่มที่ 2 ได้รับทรีทเมนต์ 2 แล้วหาค่าเฉลี่ยของตัวอย่างทั้ง 2 กลุ่มนี้ นำมาเปรียบเทียบกัน สถิติทดสอบที่ใช้คือ

$$\frac{\bar{d}}{s_d/\sqrt{n}}$$

ที่มี df = n - 1 โดยที่ d คือ ผลต่างของคะแนนแต่ละคู่ และ n คือ จำนวนคู่

เมื่อ $\bar{d} = \Sigma d/n$

$$sd = \sqrt{\Sigma (d - \bar{d})^2}$$

3.4 การใช้คำสั่ง Paired Samples T Test

ตัวอย่างเช่น ให้นักเรียน 11 คน ทำแบบทดสอบคณิตศาสตร์ชุดหนึ่ง หลังจากนั้นทำการสอนพิเศษเป็นเวลา 3 สัปดาห์ แล้วให้ทำแบบทดสอบอีกชุดหนึ่งที่มีความยากเท่า ๆ กันกับชุดแรก คะแนนที่ได้จากการสอบแต่ละครั้ง และความแตกต่างของคะแนนสอบทั้ง 2 ครั้งอยู่ในตารางที่ 1 อยากทราบว่าค่าเฉลี่ยของคะแนนสอบครั้งที่สองดีขึ้นอย่างมีนัยสำคัญที่ระดับ .05 หรือไม่

ตารางที่ 5.1 คะแนนสอบวิชาคณิตศาสตร์ 2 ครั้ง (เต็ม 90 คะแนน)

นักเรียนคนที่	1	2	3	4	5	6	7	8	9	10	11
First : การสอบครั้งแรก (1) Second : การสอบครั้งที่สอง (2)	45 53	61 67	33 47	29 34	21 31	47 49	53 62	32 51	37 48	25 29	81 86
(2) – (1)	8	6	14	5	10	2	9	19	11	4	5

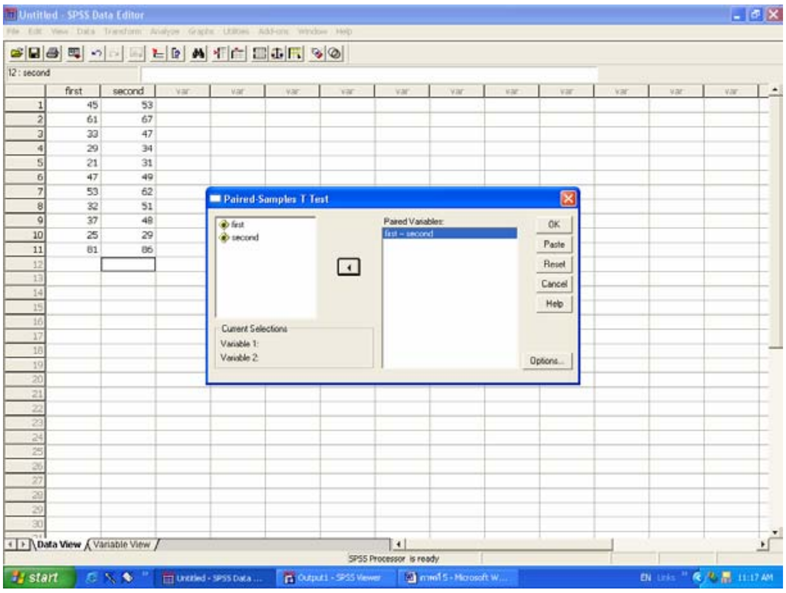
ข้อมูลคะแนนสอบวิชาคณิตศาสตร์ 2 ครั้ง อยู่ในแฟ้มข้อมูล DataTest6.sav ขั้นตอนการใช้คำสั่ง Paired – Samples T Test คือ

1. ไปที่เมนูบาร์ คลิกที่ Analyze , Compare Means, Paired – Samples T Test... จะได้หน้าต่าง Paired – Samples T Test

2. ในหน้าต่าง Paired – Samples T Test ในช่องซ้ายมือ คลิกที่ตัวแปร First และ second แล้วคลิกที่หัวลูกศร ➢ หน้าช่อง Paired Variables : ตัวแปร first และ second จะย้ายไปอยู่ในช่อง Paired Variables : ดังภาพที่ 5.5

ที่ปุ่ม Options... จะมีคำสั่งต่าง ๆ เหมือนกับการใช้คำสั่ง One – Sample T Test ในหัวข้อ 2.2 ทำเช่นเดียวกับที่อธิบายแล้ว

แล้วคลิกปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 5.6



ภาพที่ 5.5

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	first	42.18	11	17.725	5.344
	second	50.64	11	16.753	5.051

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	first & second	11	.961	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	first - second	-8.455	4.927	1.485	-11.764	-5.145	-5.692	10	.000

ภาพที่ 5.6

จากภาพผลลัพธ์ มีจำนวนตัวอย่างที่นำมาวิเคราะห์จำนวนทั้งหมด 11 คู่ มีคะแนนเฉลี่ยในการสอบครั้งแรกเท่ากับ 42.18 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 17.725 มีคะแนนเฉลี่ยในการสอบครั้งที่สองเท่ากับ 50.64 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 16.753 ค่าสัมประสิทธิ์ สหสัมพันธ์ (correlation) ระหว่างคะแนนการสอบครั้งแรกกับครั้งที่สองเท่ากับ .961 หมายความว่ามีความสัมพันธ์กันสูงมากอย่างมีนัยสำคัญ ดูที่ค่า Sig. เท่ากับ .000 ซึ่งน้อยกว่าระดับนัยสำคัญ $\alpha = .05$ ที่กำหนด

ผลการทดสอบ t – test สำหรับประชากร 2 กลุ่มที่ไม่เป็นอิสระกัน ทำการทดสอบสมมติฐานแบบทางเดียวคือ $H_0 : \mu_{first} \geq \mu_{second}$ คู่กับ $H_1 : \mu_{first} < \mu_{second}$ ซึ่งมีเขตวิกฤตอยู่ทางด้านซ้าย แล้วดูค่าสถิติ t = -5.692 ซึ่งมีค่าเป็นลบอยู่ทางด้านซ้าย ดังนั้น P-value จึงเริ่มต้นจากหางด้านซ้ายของพื้นที่ใต้โค้งไปถึง ณ ตำแหน่งของค่า t ซึ่งอยู่ด้านซ้าย จึงได้ค่า P-value เท่ากับ Sig.(2-tailed)/2 = .000/2 = .000 ซึ่งน้อยกว่า $\alpha = .05$ สรุปได้ว่าปฏิเสธ H_0 นั่นคือ คะแนนการสอบครั้งแรกน้อยกว่าครั้งที่สอง

3.5 ตัวอย่างการเปรียบเทียบระดับความคิดเห็นของผู้เข้ารับการอบรมก่อนการอบรมและหลังการอบรม

ตัวอย่างเช่น การประเมินผลการอบรมเกี่ยวกับคุณธรรม จริยธรรม และมารยาทไทย ของคณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ในปีการศึกษา 2546 กลุ่มตัวอย่างคือ นิสิตชั้นปีที่ 1 ที่เข้ารับการอบรม จำนวน 140 คน ทำการประเมินการอบรม โดยการเปรียบเทียบระดับความคิดเห็นของผู้เข้ารับการอบรมในประเด็นต่าง ๆ 10 ประเด็น ระหว่างก่อนการอบรมกับหลังการอบรม ซึ่งผู้จัดการอบรมเชื่อว่าการอบรมครั้งนี้จะสามารถกระตุ้นเตือน

หรือทำให้มีจิตสำนึกเกี่ยวกับคุณธรรม จริยธรรม และมารยาทไทยของผู้เข้ารับการอบรมได้ดีขึ้น นั่นคือ ระดับความคิดเห็นหลังการอบรมน่าจะสูงกว่าก่อนการอบรม วิธีการคือให้ผู้เข้ารับการอบรมตอบแบบสอบถามที่มีข้อความให้ผู้ตอบแสดงความคิดเห็นในประเด็นต่าง ๆ 10 ประเด็น แต่ละประเด็นมีข้อความ 4 ข้อ ในที่นี้จะนำมาเป็นตัวอย่าง 1 ประเด็นคือ เรื่องสมาธิภาวนา มีข้อคำถาม 4 ข้อ ถามเกี่ยวกับเรื่องสมาธิภาวนา การตอบข้อคำถามให้ผู้ตอบเลือกตอบระดับความคิดเห็น 5 ระดับคือ (มากที่สุด, มาก, ปานกลาง, น้อย, น้อยที่สุด) เพื่อความสะดวกในการคำนวณ จึงแปลงระดับความคิดเห็นให้เป็นคะแนนคือ

มากที่สุด	=	5 คะแนน
มาก	=	4 คะแนน
ปานกลาง	=	3 คะแนน
น้อย	=	2 คะแนน
น้อยที่สุด	=	1 คะแนน

ผู้ประเมินต้องการเปรียบเทียบระดับความคิดเห็นในเรื่องสมาธิภาวนาระหว่างก่อนการอบรมกับหลังการอบรม เรื่องสมาธิภาวนาซึ่งประกอบด้วยข้อคำถาม 4 ข้อ โดยกำหนดชื่อตัวแปรสำหรับกลุ่มก่อนการอบรมคือ it1.1, it1.2, it1.3 และ it1.4 และสำหรับกลุ่มหลังการอบรมคือ item1.1, item1.2, item1.3 และ item1.4 ข้อมูลอยู่ในแฟ้มข้อมูล Eval.sav มีรูปแบบของข้อมูลดังนี้

ตารางที่ 5.2 รูปแบบข้อมูลความคิดเห็นของข้อคำถาม it1.1, it1.2, it1.3, it1.4 และ item1.1, item1.2, item1.3, item1.4 ของนิสิต 140 คน

id	major	it1.1	it1.2	it1.3	it1.4	item1.1	item1.2	item1.3	item1.4	id	major	it1.1	it1.2	it1.3	it1.4	item1.1	item1.2	item1.3	item1.4
1	2	3	2	1	4	4	3	3	5	71	3	3	3	1	2	4	4	4	4
2	2	4	4	3	4	4	4	4	4	72	1	3	2	1	3	4	4	3	4
3	1	2	2	3	2	3	3	4	4	73	1	3	3	2	4	3	3	2	4
4	2	3	2	1	2	4	3	3	4	74	1	3	3	2	2	4	4	3	3
5	2	3	2	1	1	4	4	4	4	75	1	3	3	3	4	4	4	4	4
6	2	3	3	3	3	5	5	5	5	76	1	3	2	1	3	4	3	3	5
7	1	3	4	1	2	4	5	2	4	77	1	3	3	2	3	4	4	3	4
8	1	3	2	1	2	4	4	2	2	78	2	2	2	3	2	4	4	4	5
9	1	3	3	3	3	4	4	4	4	79	2	3	3	2	4	4	4	3	5
10	1	3	3	3	3	4	4	3	3	80	3	3	2	2	2	4	3	3	3

11	1	3	2	2	3	4	3	2	4	81	1	3	3	3	2	4	4	4	3
12	1	3	3	3	3	4	4	4	4	82	1	3	2	1	3	3	3	1	2
13	1	3	3	1	2	4	4	3	4	83	1	4	4	4	4	4	4	4	4
14	2	2	3	2	3	3	4	3	3	84	1	3	3	2	4	4	4	3	4
15	2	3	3	3	3	3	3	4	4	85	1	3	3	3	3	4	4	4	4
16	2	3	3	2	3	4	4	4	5	86	1	3	3	2	2	4	5	3	5
17	2	4	5	4	5	0	0	0	0	87	3	5	4	3	0	5	5	4	0
18	2	5	4	5	4	0	0	0	0	88	1	3	2	3	3	4	3	3	4
19	2	3	3	3	3	4	4	4	4	89	2	3	4	4	3	4	5	5	5
20	3	5	4	5	5	0	0	0	0	90	3	5	5	1	1	5	5	1	1
21	3	2	2	1	2	3	3	2	3	91	1	4	3	3	4	4	4	4	5
22	2	3	2	1	3	4	4	5	4	92	1	2	2	2	3	4	4	4	4
23	1	4	3	3	4	5	4	4	4	93	1	4	4	3	4	0	0	0	0
25	1	2	3	2	3	5	5	5	5	95	2	3	3	2	3	4	4	4	5
id	major	it1.1	it1.2	it1.3	it1.4	item1.1	item1.2	item1.3	item1.4	id	major	it1.1	it1.2	it1.3	it1.4	item1.1	item1.2	item1.3	item1.4
24	1	2	3	2	2	4	4	0	4	94	2	3	4	2	3	4	4	4	4
26	2	3	3	3	3	4	4	4	4	96	2	5	4	4	4	0	0	0	0
27	1	3	3	2	3	4	4	3	4	97	2	3	3	2	4	4	4	3	4
28	2	2	3	1	3	4	5	4	4	98	1	3	2	1	0	4	4	4	4
29	1	4	3	3	4	5	4	4	5	99	1	3	3	1	2	4	4	3	4
30	2	2	2	2	2	4	4	4	4	100	3	3	3	2	3	4	4	3	4
31	1	3	3	2	3	4	4	4	4	101	2	2	3	2	0	4	4	3	4
32	2	3	4	2	3	4	5	4	5	102	2	4	4	3	4	5	5	3	5
33	1	2	2	1	2	3	4	3	4	103	2	3	2	2	4	4	4	4	4
34	1	4	3	1	1	4	4	2	2	104	2	4	3	3	4	5	5	5	5
35	2	2	3	2	2	4	4	3	4	105	2	3	4	2	5	4	5	3	5
36	1	3	3	1	3	4	4	4	4	106	2	3	3	2	3	4	4	2	4
37	1	2	2	1	1	4	4	4	4	107	2	3	3	2	3	4	4	4	3
38	1	3	3	2	3	4	4	3	4	108	2	2	2	2	2	4	5	5	5
39	1	2	2	1	3	3	2	2	4	109	2	3	3	2	2	4	4	3	4
40	1	2	3	2	3	3	4	3	4	110	1	3	3	4	4	4	4	4	5
41	1	3	3	3	3	4	4	4	4	111	1	3	3	2	3	4	4	3	4
42	1	4	3	1	4	0	0	0	0	112	1	4	4	4	4	5	5	5	5
43	1	2	2	2	2	3	3	4	4	113	1	3	3	2	3	3	3	2	3
44	1	3	3	2	3	4	4	4	4	114	2	3	3	3	3	3	3	3	3
45	1	3	3	2	5	0	0	0	0	115	1	4	4	4	4	4	4	4	4
46	2	2	3	2	3	4	4	5	5	116	1	2	3	2	3	2	3	2	3
47	2	3	3	3	3	4	4	4	4	117	1	3	3	2	3	3	3	2	3
48	1	3	3	1	3	4	4	2	3	118	1	2	2	2	3	2	2	2	3
49	1	2	2	2	3	4	3	4	5	119	1	3	3	1	3	3	3	1	3
50	2	3	3	2	2	4	4	3	4	120	2	2	3	3	3	2	3	3	3
51	3	4	4	2	4	5	5	4	5	121	2	3	3	3	3	3	3	3	3
52	3	2	3	1	1	4	4	4	5	122	2	3	4	2	2	3	4	2	2
53	3	3	3	1	2	4	4	4	3	123	1	3	3	3	3	3	3	3	3
54	3	3	3	2	2	3	3	3	3	124	1	3	3	1	2	3	3	1	2
55	1	2	2	3	3	3	3	3	4	125	1	3	3	2	3	3	3	2	3

56	1	2	2	3	2	3	3	4	3	126	3	3	2	2	2	3	2	2	2
57	1	3	3	3	3	5	5	2	3	127	3	3	2	2	3	3	2	2	3
58	2	3	3	3	3	5	5	4	3	128	2	2	3	2	3	2	3	2	3
59	2	2	3	3	3	4	4	5	5	129	2	3	3	2	3	3	3	2	3
60	2	3	3	3	3	4	4	4	4	130	1	3	3	2	2	3	3	2	2
61	2	3	3	2	3	4	4	4	4	131	2	3	3	3	2	3	3	3	2
62	2	2	3	2	2	4	4	4	4	132	2	3	3	2	4	3	3	2	4
63	1	2	2	1	2	4	4	3	5	133	2	3	3	3	3	3	3	3	3
64	1	3	2	2	3	4	4	4	4	134	3	3	3	3	3	3	3	3	3
65	2	3	3	3	3	4	4	5	5	135	1	2	2	1	3	2	2	1	3
66	1	3	2	2	3	4	3	4	4	136	3	3	3	3	3	3	3	3	3
67	1	2	2	2	2	3	3	3	3	137	3	3	3	2	2	3	3	2	2
68	1	2	3	2	2	4	4	4	4	138	3	3	4	2	5	3	4	2	5
69	2	4	5	4	5	5	5	5	5	139	3	2	2	3	1	2	2	3	1

มีขั้นตอนการคำนวณคือ

1. สร้างตัวแปรใหม่จากการคำนวณค่าเฉลี่ยของการตอบข้อคำถาม 4 ข้อ ของกลุ่มก่อนการอบรมและกลุ่มหลังการอบรม โดยใช้คำสั่ง **Transform, Compute** ทำให้ได้ค่าเฉลี่ยของระดับความคิดเห็นเรื่องสมาธิภาวนาของตัวอย่างแต่ละคน

2. เปรียบเทียบค่าเฉลี่ยของระดับความคิดเห็นเรื่องสมาธิภาวนาของกลุ่มก่อนการอบรมกับกลุ่มหลังการอบรม เนื่องจากกลุ่มก่อนการอบรมและกลุ่มหลังการอบรมเป็นกลุ่มเดียวกันจึงเป็นการเปรียบเทียบกลุ่มตัวอย่าง 2 กลุ่มที่ไม่เป็นอิสระกัน โดยใช้คำสั่ง **Paired Samples T Test**

การใช้โปรแกรม SPSS ช่วยในการคำนวณตามขั้นตอนข้างต้นนี้คือ

1. สร้างตัวแปรใหม่ จากการคำนวณค่าเฉลี่ยของการตอบข้อคำถาม 4 ข้อ เรื่องสมาธิภาวนา

1.1 ไปที่เมนูบาร์ คลิกที่ Transform, Compute จะได้หน้าต่าง Compute Variable

1.2 ในหน้าต่าง Compute Variable

1.2.1 คำนวนค่าเฉลี่ยของตัวแปร it1.1 ถึง it1.4 ของกลุ่มก่อนการอบรม

ในช่อง Target Variable : พิมพ์ชื่อตัวแปรใหม่ m1

ในช่อง Numeric Expression : พิมพ์สูตรการคำนวณหาค่าเฉลี่ยคือ MEAN (it1.1 to it1.4)

แล้วคลิกปุ่ม OK

1.2.2 คำนวณค่าเฉลี่ยของตัวแปร item1.1 ถึง item1.4 ของกลุ่มหลังการ

อบรมในช่อง Target Variable : พิมพ์ชื่อตัวแปรใหม่ m2

ในช่อง Numeric Expression : พิมพ์สูตรการคำนวรหาค่าเฉลี่ยคือ MEAN (item1.1 to item1.4)

แล้วคลิกปุ่ม OK

2. เปรียบเทียบ m1 กับ m2

- ไปที่เมนูบาร์ คลิกที่ Analyze, Compare Means, Paired – Samples test ... จะได้หน้าต่าง Paired – Samples T Test

- ในหน้าต่าง Paired – Samples T Test

ในช่องซ้ายมือ คลิกที่ตัวแปร m1 และ m2 ให้ย้ายไปอยู่ในช่อง Paired Variables :

คลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 5.7

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	m1	2.7375	140	.61188	.05171
	m2	3.4518	140	1.03704	.08765

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	m1 & m2	140	-.179	.035

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	m1 - m2	-.71429	1.29490	.10944	-.93067	-.49790	-6.527	139	.000

ภาพที่ 5.7

จากภาพผลลัพธ์ มีจำนวนตัวอย่างที่นำมาวิเคราะห์จำนวนทั้งหมด 140 คู่ กลุ่มก่อนได้รับการอบรมมีระดับความคิดเห็นเฉลี่ยเท่ากับ 2.7375 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.61188 กลุ่มหลังได้รับการอบรมมีระดับความคิดเห็นเฉลี่ยเท่ากับ 3.4518 ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1.03704 ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างระดับความคิดเห็นของกลุ่มก่อนได้รับการอบรมและกลุ่มหลังได้รับการอบรม เท่ากับ -0.179 ดูที่ค่า Sig. เท่ากับ .035 ซึ่งน้อยกว่าระดับนัยสำคัญ α = .05 ที่กำหนด หมายความว่าคะแนนก่อนการอบรมและหลังการอบรมมีความสัมพันธ์กันแบบผกผันอย่างมีนัยสำคัญ

ผลการทดสอบ t-test สำหรับประชากร 2 กลุ่มที่ไม่เป็นอิสระกัน ทำการทดสอบสมมติฐานแบบทางเดียวคือ $H_0 : \mu_{\text{ก่อนอบรม}} \geq \mu_{\text{หลังอบรม}}$ คู่กับ $H_1 : \mu_{\text{ก่อนอบรม}} < \mu_{\text{หลังอบรม}}$ ซึ่งมีเขตวิกฤติอยู่ทางด้านซ้าย แล้วดูค่าสถิติ t = -6.527 ที่ df = 139 ค่าสถิติ t มีค่าเป็นลบอยู่ทางด้านซ้าย ดังนั้น P-value จึงเริ่มต้นจากหางด้านซ้ายของพื้นที่ใต้โค้งไปถึง ณ ตำแหน่งของค่า t ซึ่งอยู่ด้านซ้ายจึงได้ค่า P-value เท่ากับ Sig. (2-tailed)/2 = .002/2 = .000 ซึ่งน้อยกว่า α = .05 สรุปได้ว่าปฏิเสธ H_0 นั่นคือ ระดับความคิดเห็นก่อนได้รับการอบรมน้อยกว่าหลังได้รับการอบรม

4. การทดสอบสมมติฐานเกี่ยวกับค่าความแปรปรวนของประชากร 1 กลุ่ม

4.1 การวัดการกระจาย (measures of variation) ของข้อมูล

ในการศึกษาเกี่ยวกับลักษณะของข้อมูลใด ๆ นอกจากการหาค่ากลางของข้อมูลแล้ว ต้องมีการวัดการกระจายรอบ ๆ ค่ากลางนั้นด้วย เพราะบางครั้งที่ข้อมูล 2 ชุด ซึ่งมีค่ากลางเหมือนกันอาจมีการกระจายของข้อมูลแตกต่างกัน การวัดค่ากลางของข้อมูลที่ใช้กันมากคือค่าเฉลี่ย $(\bar{x})$ เป็นค่าเฉลี่ยของกลุ่มตัวอย่าง และค่าเบี่ยงเบน (deviation) จากค่าเฉลี่ยคือ $x - \bar{x}$ ซึ่งผลรวมของค่าเบี่ยงเบนของข้อมูลทั้งหมดเท่ากับ 0 นั่นคือ $\sum(x_i - \bar{x}) = 0$ ในการวัดการกระจายของข้อมูลเราจึงต้องยกกำลังสองค่าเบี่ยงเบนเหล่านี้ทุกตัวรวมกันทั้งหมดแล้วหารด้วยจำนวนข้อมูลทั้งหมดลบ 1 เรียกว่า ความแปรปรวนของกลุ่มตัวอย่าง (sample variance) มีสูตร คือ

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

เมื่อ n คือ ขนาดของตัวอย่าง

และรากที่สองของความแปรปรวนของกลุ่มตัวอย่างคือ ค่าความเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง (sample standard deviation) เขียนแทนด้วย S

สำหรับความแปรปรวนของประชากร (population variance) คำนวณได้จากสูตร

$$\sigma^2 = \sum x_i^2 f(x_i) - \mu^2$$

เมื่อ $f(x) = P(X = x)$ เรียกว่าฟังก์ชันการแจกแจงความน่าจะเป็น (probability distribution function) หรือเขียนย่อ ๆ คือ p.d.f. ของ X

และ σ คือ ค่าความเบี่ยงเบนมาตรฐานของประชากร (population standard deviation) ซึ่งมีหน่วยเหมือนหน่วยของ X ตัวอย่างเช่น ถ้า X คือรายได้มีหน่วยเป็นบาท σ ก็จะมีหน่วยเป็นบาท เรามักวัดการกระจายของข้อมูลด้วย σ มากกว่าการวัดด้วย σ^2

4.2 การทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ σ^2

นอกจากการสรุปอ้างอิงเกี่ยวกับค่าเฉลี่ยของประชากรแล้ว เรายังต้องการสรุปอ้างอิงเกี่ยวกับการกระจายของประชากรด้วย เพราะมันสามารถเป็นตัวชี้วัดความเที่ยงตรง (reliability) ของข้อมูล คือ การมีรูปแบบการกระจายของข้อมูลแบบเดียวกัน ซึ่งสามารถแสดงมาตรฐานของคุณภาพของกระบวนการผลิตสินค้าได้ ในการควบคุมคุณภาพของสินค้า เรามักจะต้องการความมั่นใจว่าค่าการกระจายของสินค้าหรือสิ่งที่สนใจจะไม่เกินค่าที่กำหนด ซึ่งผลิตมาจากเครื่องจักรหลาย ๆ ตัว ในการสรุปอ้างอิงเกี่ยวกับค่าความเบี่ยงเบนมาตรฐานของประชากร (σ) มีข้อตกลงเบื้องต้นคือ ประชากรมีการแจกแจงแบบปกติ

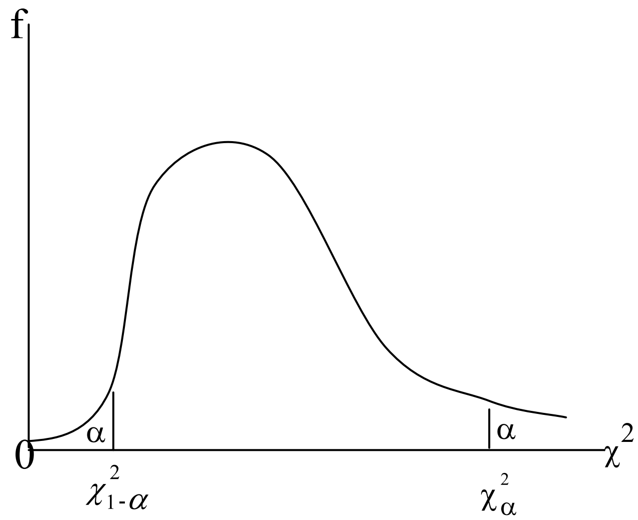
สมมติฐานที่ต้องการทดสอบคือ $H_0 : \sigma^2 = \sigma_0^2$

สถิติที่ใช้ทดสอบคือ χ^2 – test มีสูตรคือ

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \quad ; \quad df = n-1$$

ถ้าต้องการทดสอบแบบทางเดียว สมมติฐานแย้ง $H_1 : \sigma^2 > \sigma_0^2$ เขตปฏิเสธของการทดสอบที่มีระดับนัยสำคัญ α คือ ค่า χ^2 ที่ได้จากการคำนวณมากกว่าหรือเท่ากับ χ^2_α ที่เปิดจากตารางที่ df = n - 1

ถ้าต้องการทดสอบแบบสองทาง สมมติฐานแย้ง $H_1 : \sigma^2 \neq \sigma_0^2$ เขตปฏิเสะของการทดสอบที่มีระดับนัยสำคัญ α คือ ค่า χ^2 ที่ได้จากการคำนวณน้อยกว่าหรือเท่ากับ $\chi^2_{-\alpha/2}$ หรือมากกว่าหรือเท่ากับ $\chi^2_{\alpha/2}$



ภาพที่ 5.8 โค้งการแจกแจงความน่าจะเป็นการแจกแจงไคสแคว

4.3 ตัวอย่างการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ σ^2

ตัวอย่างเช่น การศึกษาเกี่ยวกับค่าความผันแปรของสินค้าที่ผลิตคือ นาฬิกาข้อมือ ผู้ผลิตอ้างว่าการทำงานของนาฬิกาข้อมือมีการแจกแจงแบบปกติ และมีค่าเบี่ยงเบนมาตรฐานเป็น 0.2

วินาที โดยสุ่มตัวอย่างนาฬิกาข้อมือมาจำนวน 10 เรือน จากผลผลิตจำนวนมากที่ผ่านขั้นตอนการตรวจสอบคุณภาพแล้ว เมื่อครบ 1 เดือน ทำการบันทึกข้อมูลเวลาของนาฬิกาข้อมือทั้ง 10 เรือน ที่เบี่ยงเบนจากนาฬิกามาตรฐาน แล้วคำนวณค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของตัวอย่างได้ $\bar{x}$ = .7 วินาที , S = .4 วินาที

สมมติฐานที่ต้องการทดสอบคือ $H_0 : \sigma^2 = .04$ คู่กับ $H_1 : \sigma^2 \neq .04$

คำนวณค่าสถิติทดสอบ

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(10-1)(.4)^2}{.04} = 36.0 \quad ; \quad df = 10-1 = 9$$

เปิดตารางเปอร์เซ็นไทล์ของการแจกแจงแบบไคสแคว ; $\chi^2_{\alpha/2}$ สำหรับ $P(\chi^2 \geq \chi^2_{\alpha/2}) = \alpha$ ที่ $\alpha = .05$ ได้ค่า $\chi^2_{.025} = 19.02$ และ $\chi^2_{.975} = 2.70$ เขตปฏิเสธ H_0 คือ $\chi^2 \leq \chi^2_{-\alpha/2}$ หรือ $\chi^2 \geq \chi^2_{\alpha/2}$ เปรียบเทียบค่า χ^2 ที่คำนวณได้กับค่าที่เปิดจากตารางพบว่า χ^2 มีค่ามากกว่า $\chi^2_{.025}$ ดังนั้นจึงสรุปได้ว่าปฏิเสธ H_0 นั่นคือ ความแปรปรวนของการทำงานของนาฬิกาข้อมือที่ผลิตได้จากโรงงานไม่เท่ากับ 0.2 (วินาที)^2

5. การทดสอบสมมติฐานเกี่ยวกับค่าความแปรปรวนของประชากร 2 กลุ่ม

5.1 การทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ $\sigma_1^2 = \sigma_2^2$

มีวัตถุประสงค์เพื่อเปรียบเทียบค่าความแปรปรวนของประชากร 2 กลุ่มที่เป็นอิสระกันว่าแตกต่างกันหรือไม่

ข้อตกลงเบื้องต้นคือ กลุ่มตัวอย่างทั้ง 2 กลุ่ม สุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ

สมมติฐานที่ต้องการทดสอบคือ $H_0 : \sigma_1^2 = \sigma_2^2$ คู่กับ $H_1 : \sigma_1^2 \neq \sigma_2^2$

สถิติที่ใช้ทดสอบคือ F-test มีสูตรคือ

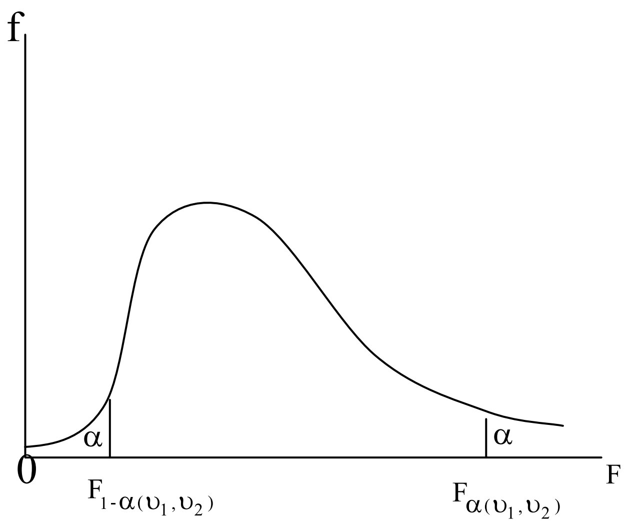
$$F = \frac{S_1^2}{S_2^2} \quad ; \; df = n_1 - 1 \text{ และ } n_2 - 1$$

เมื่อ S_1^2 คือ ค่าความแปรปรวนของตัวอย่างกลุ่มที่ 1

S_2^2 คือ ค่าความแปรปรวนของตัวอย่างกลุ่มที่ 2

n_1 , n_2 คือ จำนวนของตัวอย่างกลุ่มที่ 1 และ 2

เขตปฏิเสธของการทดสอบที่มีระดับนัยสำคัญ α คือ $F \le F_{1-\alpha/2}$ หรือ $F \ge F_{\alpha/2}$



ภาพที่ 5.9 โค้งการแจกแจงความน่าจะเป็นของการแจกแจงไคสแควร์

5.2 ตัวอย่างการทดสอบสมมติฐานเกี่ยวกับพารามิเตอร์ $\sigma_1^2 = \sigma_2^2$

ตัวอย่างเช่น การศึกษาเกี่ยวกับค่าความผันแปรของผลผลิตนมจากโคนมในฟาร์มของเกษตรกร 2 แห่ง อยากทราบว่าน้ำหนักผลผลิตนมของโคนมในฟาร์มทั้ง 2 แห่งมีค่าเบี่ยงเบน

เท่ากันหรือไม่ โดยสุ่มตัวอย่างโคนมจากฟาร์มแรก จำนวน 13 ตัว และฟาร์มที่สอง จำนวน 12 ตัว แล้วเก็บข้อมูลน้ำหนักนมต่อวันของโคนมแต่ละตัวในระยะเวลา 2 สัปดาห์ แล้วหาค่าเฉลี่ยเป็นน้ำหนักนมต่อวัน ได้ข้อมูลดังตาราง

ตารางที่ 5.3 น้ำหนักนมต่อวัน (กิโลกรัม) ของโคนมจากฟาร์ม 2 แห่ง

ฟาร์มที่ 1	44	44	56	45	47	38	58	53	49	35	46	30	41
ฟาร์มที่ 2	35	47	55	29	40	39	32	41	42	57	51	39	

สมมติฐานที่ต้องการทดสอบคือ $H_0 : \sigma_1^2 = \sigma_2^2$ คู่กับ $H_1 : \sigma_1^2 \neq \sigma_2^2$

ในการทดสอบสมมติฐานข้างต้นสามารถใช้โปรแกรม SPSS ช่วยในการคำนวณได้โดยใช้คำสั่ง One - Way ANOVA ซึ่งเป็นการวิเคราะห์ความแปรปรวนแบบทางเดียว

5.3 การใช้คำสั่ง One - Way ANOVA

ข้อมูลน้ำหนักนมต่อวันของโคนมจากฟาร์ม 2 แห่ง อยู่ในแฟ้มข้อมูล Datacow.sav ที่มีตัวแปร farm (1 = ฟาร์ม 1 , 2 = ฟาร์ม 2) และตัวแปร yield แทนน้ำหนักนมต่อวัน การทดสอบสมมติฐานดังกล่าวข้างต้นมีขั้นตอนการใช้คำสั่งดังนี้

1. ไปที่เมนูบาร์ คลิกที่ Analyze, Compare Means, One – Way ANOVA… จะได้หน้าต่าง One – Way ANOVA

2. ในหน้าต่าง One – Way ANOVA คลิกที่ตัวแปรตาม yield ย้ายเข้าไปอยู่ในช่อง Dependent List : และคลิกที่ตัวแปร farm ย้ายของไปอยู่ในช่อง Factor :

คลิกที่ปุ่ม Options… จะได้หน้าต่าง One – Way ANOVA : Options

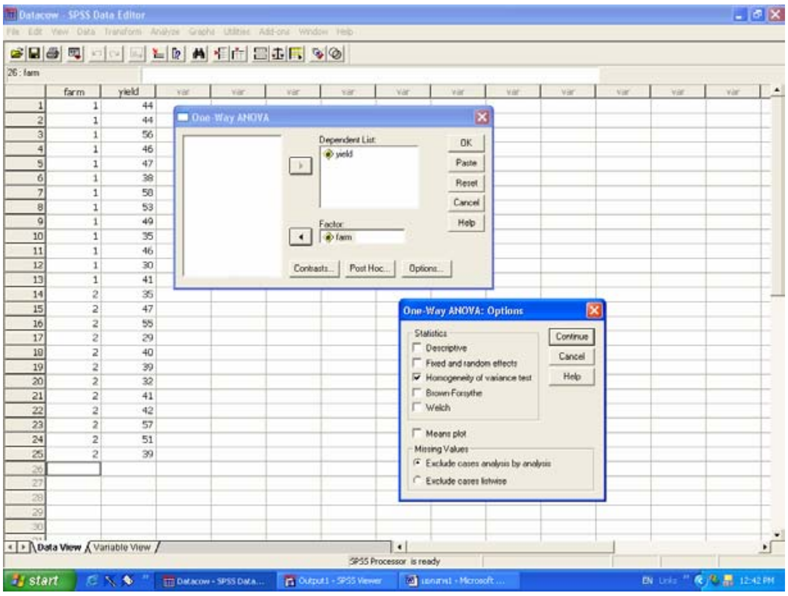
3. ในหน้าต่าง One – Way ANOVA : Options

คลิกที่ ☐ Homogeneity of variance test เพื่อทดสอบสมมติฐาน $H_0 : \sigma_1^2 = \sigma_2^2$ ดังภาพที่ 5.10

คลิกที่ปุ่ม Continue หน้าต่างนี้จะถูกปิดไป

4. ในหน้าต่าง One – Way ANOVA

คลิกที่ปุ่ม OK จะได้ผลลัพธ์ดังภาพที่ 5.11



ภาพที่ 5.10

Oneway

Test of Homogeneity of Variances

yield

Levene Statistic	df1	df2	Sig.
.176	1	23	.678

ANOVA

yield

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	52.618	1	52.618	.753	.395
Within Groups	1607.942	23	69.911		
Total	1660.560	24			

ภาพที่ 5.11

จากภาพผลลัพธ์ในตาราง Test of Homogeneity of Variances ดูค่าสถิติ Levene Statistic เท่ากับ .176 ที่ $df_1 = 1$, $df_2 = 23$ และค่า Sig. เท่ากับ .678 ซึ่งมากกว่าระดับนัยสำคัญที่

กำหนด (α = .05) จึงสรุปได้ว่ายอมรับ $H_0 : \sigma_1^2 = \sigma_2^2$ นั่นคือ น้ำหนักของผลผลิตนมของโคนมในฟาร์มทั้ง 2 แห่ง มีความแปรปรวนเท่ากัน