

การวิเคราะห์ข้อมูลจากแผนแบบการทดลองสุ่มสมบูรณ์ กรณีความแปรปรวนไม่เท่ากันด้วยโปรแกรมสำเร็จรูปทางสถิติ

Analysis of Data from Completely Randomized Designs Using Statistical Packages with Heterogeneity of Error Variance

กมล นุชบา

ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ปทุมธานี 12121

kamon@mathstat.sci.tu.ac.th

1. คำนำ

การวิเคราะห์ข้อมูลจากแผนแบบการทดลองสุ่มสมบูรณ์หรือ CRD (Completely Randomized Designs) นั้นใช้วิธีการทางสถิติที่เรียกว่าการวิเคราะห์ความแปรปรวนโดยอาศัยสถิติเอฟเพื่อทดสอบว่าค่าเฉลี่ยประชากรของแต่ละทรีทเมนต์มีค่าเท่ากันหรือไม่ หรือทดสอบว่าอิทธิพลของทรีทเมนต์มีนัยสำคัญหรือไม่

อย่างไรก็ตาม โดยทั่วไปนักวิจัยมักจะไม่นำข้อสังเกตเกี่ยวกับความสนใจกับการตรวจสอบว่าข้อมูลที่น่าไปวิเคราะห์ด้วยการวิเคราะห์ความแปรปรวนนั้นเป็นไปตามข้อตกลงเบื้องต้นของการใช้สถิติเอฟหรือไม่ และในการนำเสนอผลการวิจัยก็มักจะไม่นำข้อกล่าวถึงวิธีการที่ใช้ในการตรวจสอบว่าข้อมูลนั้นสามารถใช้การวิเคราะห์ความแปรปรวนได้อย่างน่าเชื่อถือหรือไม่ ทำให้ผู้อ่านหรือผู้ที่ทำงานวิจัยในลักษณะเดียวกันสับสนหรือมองข้ามความสำคัญของการตรวจสอบข้อตกลงเบื้องต้นของวิธีการทางสถิติ

2. ข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวน

ข้อตกลงเบื้องต้นคือเงื่อนไขที่กำหนดให้กับความคลาดเคลื่อนสุ่มที่อยู่ในตัวแบบที่ใช้อธิบายค่าสังเกตแต่ละค่า (y_{ij}) ที่ได้จากการทดลอง สำหรับแผนแบบการทดลอง

แบบสุ่มสมบูรณ์ที่อิทธิพลของทรีทเมนต์เป็นค่าคงที่มีตัวแบบทางสถิติคือ

$$y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

โดยที่ $i = 1, 2, 3, \dots, t$ ทรีทเมนต์ และ $j = 1, 2, 3, \dots, r_i$ ซ้ำ ในที่นี้ μ แทนค่าเฉลี่ยประชากรซึ่งเป็นค่าร่วมกันของค่าสังเกตทุกค่า μ_i แทนค่าเฉลี่ยประชากรของทรีทเมนต์ที่ i และ τ_i แทนอิทธิพลที่แท้จริงของทรีทเมนต์ที่ i (มีค่าเท่ากับ $\mu_i - \mu$) ส่วน ε_{ij} แทนความคลาดเคลื่อนสุ่ม ข้อตกลงเบื้องต้นที่ระบุให้กับ ε_{ij} นั้นได้แก่ ความคลาดเคลื่อนสุ่มนี้มาจากประชากรเดียวกันที่มีการแจกแจงแบบปกติโดยมีค่าเฉลี่ยศูนย์และความแปรปรวน σ^2 เท่ากันในทุกทรีทเมนต์และแต่ละตัวต่างก็เป็นอิสระซึ่งกันและกัน [1]

3. ผลของการไม่เป็นไปตามข้อตกลงเบื้องต้น

ผลของการไม่เป็นไปตามข้อตกลงเบื้องต้นของความคลาดเคลื่อนสุ่มมีกล่าวไว้ในหนังสือสถิติหลายเล่ม เช่น Kutner, et al. [2], Oehlert, G. R. [3] และ Kuehl, R. O. [1] และสรุปสาระสำคัญได้ดังนี้

สำหรับตัวแบบการวิเคราะห์ความแปรปรวนที่มีอิทธิพลของทรีทเมนต์เป็นค่าคงที่ การบ่งชี้ไปจากการแจกแจงแบบปกติถือว่าไม่ใช่เรื่องที่สำคัญเท่าไรกับการบ่งชี้ไปนั้น ไม่ถึงขั้นรุนแรง เนื่องจากสถิติเอฟมีความ

แกร่ง (robust) ต่อการย้ายเบนไปจากการแจกแจงแบบปกติ ดังนั้นจึงมีผลเพียงเล็กน้อยต่อระดับนัยสำคัญและกำลังการทดสอบ ([1] และ [2]) นอกจากนั้นความโค้งของการแจกแจงของความคลาดเคลื่อนมีผลต่อการทดสอบสมมติฐานมากกว่าความเบ้ ([2] และ [3]) แต่สำหรับตัวแบบที่อิทธิพลของทรีเมนต์เป็นตัวแปรสุ่ม การย้ายเบนไปจากการแจกแจงแบบปกติจะก่อให้เกิดปัญหาที่ร้ายแรง [2]

สำหรับในกรณีที่มีความแปรปรวนของความคลาดเคลื่อนสุ่มมีค่าไม่เท่ากันนั้นมีผลต่อการทดสอบด้วยสถิติเอฟบ้างถ้าจำนวนซ้ำในแต่ละทรีเมนต์เท่ากันสำหรับตัวแบบที่อิทธิพลของทรีเมนต์เป็นค่าคงที่แต่จะมีผลอย่างรุนแรงสำหรับตัวแบบที่อิทธิพลของทรีเมนต์เป็นตัวแปรสุ่มไม่ว่าจำนวนซ้ำจะเท่ากันในแต่ละทรีเมนต์หรือไม่ก็ตาม [2] อย่างไรก็ตาม Kuehl [1] กล่าวถึงผลกระทบที่มีต่อระดับนัยสำคัญของการทดสอบด้วยสถิติเอฟโดยอ้างอิงจากงานของ Box ในปี ค.ศ. 1954 ว่า ถ้าทำการทดสอบด้วยสถิติเอฟที่ระดับนัยสำคัญจริง .05 ถ้าอัตราส่วนของความแปรปรวนที่มีค่าน้อยสุดต่อความแปรปรวนที่มีค่ามากที่สุดในการทดลองนั้นคือ 1:3 แล้ว ระดับนัยสำคัญที่คำนวณได้จะมีค่าระหว่าง .056 ถึง .074 ถ้าจำนวนซ้ำในแต่ละทรีเมนต์เท่ากัน แต่จะมีค่าระหว่าง .013 ถึง .140 ถ้าจำนวนซ้ำในแต่ละทรีเมนต์ไม่เท่ากัน

ถ้าความคลาดเคลื่อนสุ่มไม่เป็นอิสระกัน จะมีผลอย่างรุนแรงต่อการอ้างอิงในการวิเคราะห์ความแปรปรวนทั้งในกรณีที่อิทธิพลของทรีเมนต์เป็นค่าคงที่และเป็นตัวแปรสุ่ม [2] กล่าวคือ ค่าคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยของทรีเมนต์จะเป็นตัวประมาณที่เอนเอียง โดยจะมีสูงเกินจริงถ้าความคลาดเคลื่อนมีความสัมพันธ์ในเชิงบวกและมีค่าเกินจริงถ้าความคลาดเคลื่อนมีความสัมพันธ์เชิงลบ ([1] และ [3]) และเนื่องจากการตรวจสอบในเรื่องนี้ทำค่อนข้างยาก จึงเป็นสิ่งสำคัญที่ควรจะมีการป้องกันตั้งแต่ขั้นต้นของการทดลอง เช่น ใช้การสุ่ม (randomization) หน่วยทดลองเพื่อรับทรีเมนต์และใช้การสุ่มลำดับของการวัดหรือบันทึกค่าสังเกต ในกรณีที่ทราบแน่ชัดว่าความคลาดเคลื่อนมีความสัมพันธ์กันจะใช้วิธีปรับปรุงตัวแบบที่

ใช้วิเคราะห์ข้อมูล เช่น ใช้ตัวแบบสำหรับแผนแบบที่มีการวัดซ้ำ (repeated measures designs) เป็นต้น [2]

4. การตรวจสอบด้วยส่วนตกค้าง (Residuals)

สิ่งที่ใช้เป็นพื้นฐานสำหรับการตรวจสอบการเป็นไปตามข้อตกลงเบื้องต้นของตัวแบบเชิงเส้นได้แก่ “ส่วนตกค้าง (residuals)” ส่วนตกค้างใช้เป็นตัวประมาณความคลาดเคลื่อนสุ่มและสำหรับในตัวของแผนแบบการทดลองสุ่มสมบูรณ์คำนวณโดย $\hat{\epsilon}_{ij} = y_{ij} - \bar{y}_i$ หรือผลต่างระหว่างค่าสังเกตกับค่าเฉลี่ยของทรีเมนต์ของค่าสังเกตนั้นๆ การตรวจสอบทำได้ทั้งการใช้การลงจุดหรือกราฟของส่วนตกค้างและการใช้สถิติในการทดสอบ

การตรวจสอบว่าความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติหรือไม่ อาจพิจารณาจากฮิสโทแกรม หรือการลงจุดความน่าจะเป็นแบบปกติ (normal probability plots) ของส่วนตกค้าง หรือใช้สถิติทดสอบ เช่น การทดสอบโคลโมโกรอฟ สเมอร์นอฟ (Kolmogorov-Smirnov) หรือที่ได้ปรับปรุงขึ้นเป็นการทดสอบลิลลี่ฟอว์ (Lilliefors), หรือการทดสอบชาฟิโร-วิลค์ (Shapiro-Wilk) เป็นต้น อย่างไรก็ตามการตรวจสอบข้อตกลงเบื้องต้นในข้อนี้ไม่จำเป็นเท่าใดนักเนื่องจากสถิติเอฟมีความแข็งแกร่งต่อการย้ายเบนไปจากความเป็นปกติตามที่ได้กล่าวมาแล้ว

การตรวจสอบว่าความแปรปรวนของความคลาดเคลื่อนสุ่มมีค่าเท่ากันหรือไม่ทำได้โดย การพิจารณาจากแผนภาพกล่องของส่วนตกค้างจำแนกตามทรีเมนต์ หรือใช้การลงจุดระหว่างส่วนตกค้างกับค่าเฉลี่ยของแต่ละทรีเมนต์ซึ่งจะทำให้ทราบว่าความแปรปรวนของความคลาดเคลื่อนมีความสัมพันธ์กับค่าเฉลี่ยดังกล่าวในลักษณะใด ส่วนการทดสอบว่าความแปรปรวนของความคลาดเคลื่อนมีค่าเท่ากันหรือไม่ อาจใช้การทดสอบเลวิน (Levene test) อย่างไรก็ตามการใช้สถิติทดสอบโดยไม่สนใจกราฟของส่วนตกค้าง จะทำให้ไม่ทราบถึงความสัมพันธ์ระหว่างความแปรปรวนของความคลาดเคลื่อนสุ่มและค่าเฉลี่ยของทรีเมนต์ ซึ่งจะช่วยใน

การหาวิธีแก้ไขให้ความคลาดเคลื่อนมีความแปรปรวนเท่ากันซึ่งจะได้กล่าวถึงต่อไป

ถ้าค่าสังเกตหรือการบันทึกข้อมูลเป็นไปตามลำดับเวลาการตรวจสอบว่าความคลาดเคลื่อนมีความสัมพันธ์กันหรือไม่นั้นสามารถใช้การลงจุดส่วนตกค้างตามลำดับเวลาที่บันทึกข้อมูล หรือใช้การทดสอบด้วยสถิติเคอร์บิน-วัตสัน (Durbin-Watson) อย่างไรก็ตามการทดลองส่วนใหญ่ไม่ได้ทำการบันทึกข้อมูลตามลำดับเวลาและใช้การสุ่มทริทเมนต์หรือลำดับการวัดค่าสังเกต ทำให้ไม่จำเป็นที่จะต้องตรวจสอบในเรื่องนี้

5. การแก้ไขให้เป็นไปตามข้อตกลงเบื้องต้น

วิธีการแก้ไขให้ข้อมูลเป็นไปตามข้อตกลงเบื้องต้นในเรื่องของความแปรปรวนของความคลาดเคลื่อนสุ่มที่มีค่าไม่เท่ากัน Kutner, et al. [2] ได้สรุปวิธีการ ดังนี้

5.1 ถ้าความคลาดเคลื่อนมีการแจกแจงแบบปกติแต่ความแปรปรวนของความคลาดเคลื่อนมีค่าแตกต่างกันให้ใช้วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก (weighted least squares)

5.2 ถ้าความคลาดเคลื่อนไม่ได้มีการแจกแจงแบบปกติและความแปรปรวนก็มีค่าแตกต่างกันให้ใช้การแปลงข้อมูล (transform) โดยอาจใช้วิธีการแปลงข้อมูลตามเกณฑ์ที่กำหนดถ้าทราบความสัมพันธ์ระหว่างความแปรปรวนและค่าเฉลี่ยหรือทราบลักษณะของข้อมูล หรือใช้วิธีการของบ็อกซ์-คอกซ์ (Box-Cox) ก็ได้

5.3 ถ้าความคลาดเคลื่อนปายเบนไปจากข้อตกลงเบื้องต้นของตัวแบบมากและไม่สามารถใช้การแปลงข้อมูลเพื่อให้ความคลาดเคลื่อนมีความแปรปรวนเท่ากันและมีการแจกแจงที่ไม่ใกล้เคียงแบบปกติ ให้ใช้การทดสอบที่ไม่อิงพารามิเตอร์ (non-parametric test)

สำหรับในกรณีที่ความคลาดเคลื่อนมีความสัมพันธ์กันวิธีการแก้ไขค่อนข้างซับซ้อน เช่น ใช้วิธีการวิเคราะห์แบบอนุกรมเวลาและสถิติเชิงปริภูมิ (spatial statistics) เป็นต้น [3]

อนึ่ง การแปลงข้อมูลให้ความแปรปรวนมีค่าเท่ากันอาจทำได้โดยการยกกำลังข้อมูลเดิม โดยประมาณ “ด้วยยกกำลัง” จากการถดถอยค่าลอการิทึมของส่วนเบี่ยงเบนมาตรฐานบนค่าลอการิทึมของค่าเฉลี่ยของแต่ละทริทเมนต์ และเมื่อได้ค่าประมาณของความชันของเส้นถดถอยคือ b จะได้ค่าประมาณของด้วยยกกำลังของข้อมูลเดิมคือ $1-b$ [1]

6. วิธีการของบราวน์-ฟอร์ซิท (Brown-Forsythe)

นอกจากการแก้ไขให้ข้อมูลเป็นไปตามข้อตกลงเบื้องต้นแล้วใช้สถิติเอฟทดสอบแล้ว วิธีการที่วิเคราะห์ข้อมูลในกรณีที่ความแปรปรวนของความคลาดเคลื่อนสุ่มมีค่าไม่คงที่ วิธีหนึ่งได้แก่ วิธีการของบราวน์-ฟอร์ซิท สถิติดังกล่าวคำนวณ โดย [3]

$$F_{BF} = \frac{\sum_{i=1}^I r_i (\bar{y}_i - \bar{y}_{..})^2}{\sum_{i=1}^I (1 - r_i / N) s_i^2}$$

เมื่อ s_i^2 คือความแปรปรวนตัวอย่างในทริทเมนต์ที่ i และ N คือจำนวนหน่วยทดลองหรือจำนวนค่าสังเกตทั้งหมด และภายใต้สมมุติฐานว่างที่เป็นจริง สถิตินี้มีการแจกแจงแบบเอฟโดยประมาณที่มีระดับขั้นความเสรีคือ $t-1$ และ v โดยที่

$$v = \frac{(\sum_{i=1}^I d_i)^2}{\sum_{i=1}^I d_i^2 / (r_i - 1)}$$

$$\text{และ } d_i = (1 - r_i / N) s_i^2$$

ให้สังเกตว่าความแตกต่างระหว่างสถิติเอฟกับ F_{BF} นั้นต่างกันที่เทอมส่วนหรือวิธีที่ใช้ประมาณค่าความแปรปรวนภายในกลุ่มนั่นเอง

7. วิธีการทดสอบของเวลช์ (Welch Test)

อีกวิธีการหนึ่งที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยโดยไม่มีข้อตกลงเบื้องต้นว่าความแปรปรวนของความคลาดเคลื่อนสุ่มต้องมีค่าเท่ากัน ได้แก่การทดสอบของเวลช์ สถิติดังกล่าวคำนวณ โดย [4]

$$F_{Welch} = \frac{\sum_{i=1}^t w_i [(\bar{y}_i - \bar{y}_{..})^2 / (t-1)]}{1 + \frac{2(t-2)}{(t^2-1)} \sum_{i=1}^t [(1 - \frac{w_i}{u}) / (r_i - 1)]}$$

เมื่อ $w_i = r_i / s_i^2$, $u = \sum_{i=1}^t w_i$, และ

$\bar{y}_{..} = \sum_{i=1}^t w_i \bar{y}_i / u$ และเมื่อสมมติฐานว่างเป็นจริง สถิติ

นี้มีการแจกแจงแบบเอฟโดยประมาณ ที่มีระดับชั้นความเสรีคือ $t-1$ และ ν โดยที่

$$\nu = \left[\frac{3}{t^2 - 1} \sum_{i=1}^t (1 - \frac{w_i}{u})^2 / (r_i - 1) \right]^{-1}$$

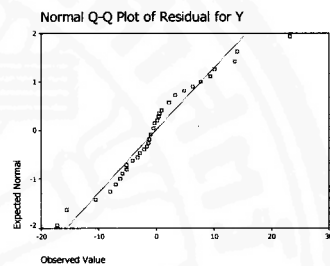
8. การใช้โปรแกรมสำเร็จรูป SPSS

พิจารณาตัวอย่างการทดสอบอายุการใช้งานของยางสนที่ใช้หุ้มสายเชื่อมในวงจรเบ็ดเสร็จหรือวงจร IC (Integrated Circuits) ที่ระดับอุณหภูมิต่างๆ ถ้าสุ่มชิ้นส่วนสายเชื่อมจำนวน 37 ชิ้น เพื่อทดสอบที่อุณหภูมิ 5 ระดับ ข้อมูลต่อไปนี้แสดงจำนวนชั่วโมงที่ทดสอบจนกว่าสายเชื่อมจะใช้งานได้

ตารางที่ 1 จำนวนชั่วโมงที่ใช้ทดสอบจนกว่าสายเชื่อมจะใช้งานได้ภายใต้อุณหภูมิระดับต่างๆ

อุณหภูมิ	ชั่วโมง			
175 °C	109.648	81.2831	100.000	83.1764
	70.7946	91.2011	75.8578	79.4328
194 °C	45.7088	51.2861	26.3027	57.5440
	45.7088	40.7380	35.4813	45.7088
213 °C	33.8844	34.6737	23.9883	20.4174
	22.3872	18.6209	18.1970	23.9883
231 °C	14.1254	16.5959	14.7911	14.4544
	16.2181	19.0546	14.7911	
250 °C	18.1970	6.7608	12.0226	10.4713
	12.3027	11.4815		

จากข้อมูลในตารางที่ 1 หลังจากคำนวณส่วนตกค้างและใช้เมนูคำสั่ง Analyze ตามด้วยเมนู Explore และเลือกให้แสดงการลงจุดความน่าจะเป็นแบบปกติ และสถิติที่ใช้ทดสอบว่าความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติหรือไม่ ผลลัพธ์ที่ได้แสดงในภาพที่ 1 จากนั้นเลือกให้แสดงแผนภาพกล่องจำแนกตามทรีทเมนต์และสถิติเฉลี่ยเพื่อทดสอบว่าความแปรปรวนของความคลาดเคลื่อนสุ่มมีค่าเท่ากันหรือไม่ ผลลัพธ์ที่ได้แสดงในภาพที่ 2

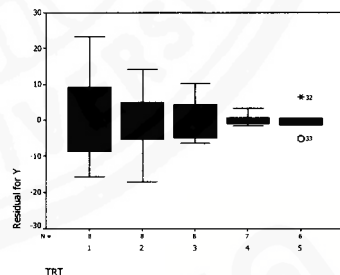


Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual for Y	.149	37	.038	.955	37	.138

a. Lilliefors Significance Correction

a. Lilliefors Significance Correction

ภาพที่ 1 ผลลัพธ์แสดงการลงจุดและสถิติที่ใช้ทดสอบว่าความคลาดเคลื่อนสุ่มมีการแจกแจงแบบปกติหรือไม่



Test of Homogeneity of Variance				
Residual for Y	Levene Statistic	df1	df2	Sig.
	Based on Mean	4.480	4	.005
	Based on Median	2.375	4	.073
	Based on Median and with adjusted df	2.375	4	.091
	Based on trimmed mean	4.196	4	.008

ภาพที่ 2 ผลลัพธ์แสดงแผนภาพกล่องของส่วนตกค้างจำแนกตามทรีทเมนต์และสถิติที่ใช้ทดสอบว่าความแปรปรวนของความคลาดเคลื่อนสุ่มมีค่าเท่ากันหรือไม่

จากการลงจุดความน่าจะเป็นแบบปกติในภาพที่ 1 การเรียงของส่วนตกค้างบายนไปจากเส้นตรงไม่มากนัก แสดงว่าความคลาดเคลื่อนสุ่มมีการแจกแจงที่บายนไปจากการแจกแจงแบบปกติไม่มากซึ่งสอดคล้องกับการทดสอบของชาปี-โรวิลล์ที่ให้ค่า p-value เท่ากับ .138 อย่างไรก็ตามหากพิจารณาผลการทดสอบของโคลโมโกรอฟ สเมอร์นอฟ พบว่ามีนัยสำคัญที่ระดับนัยสำคัญ .05 (p-value เท่ากับ .038) แต่สิ่งที่สำคัญที่ควรให้ความสำคัญมากกว่าการทดสอบการแจกแจงแบบปกติคือแผนภาพกล่องของส่วนตกค้างจำแนกตามทรีทเมนต์ชี้ให้เห็นว่าความแปรปรวนของความคลาดเคลื่อนมีมากในทรีทเมนต์ที่ 1 และ 2 และมากกว่าในทรีทเมนต์ที่ 4 และ 5 ซึ่งสอดคล้องกับผลการทดสอบด้วยสถิติเลวินที่คำนวณโดยใช้ค่าเฉลี่ยและใช้ค่าเฉลี่ยที่ตัดแต่ง (trimmed mean) เนื่องจาก p-value ของการทดสอบเท่ากับ .005 และ .008 ตามลำดับ ทำให้ผลการทดสอบด้วยสถิติเอฟตามปกติในตารางวิเคราะห์ความแปรปรวนในภาพที่ 3 มีความน่าเชื่อถือน้อย

ANOVA

Y	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28027.85	4	7006.964	101.810	.000
Within Groups	2202.365	32	68.824		
Total	30230.22	36			

ภาพที่ 3 การทดสอบเอฟในตารางวิเคราะห์ความแปรปรวน

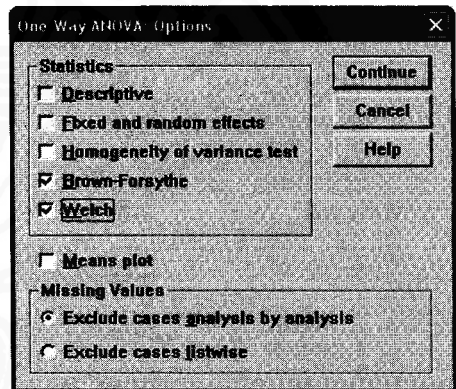
อย่างไรก็ตามในโปรแกรม SPSS รุ่น 11.0 ขึ้นไปในเมนูคำสั่ง One-Way ANOVA มีทางเลือกให้แสดงสถิติทดสอบค่าเฉลี่ยของทรีทเมนต์เมื่อความแปรปรวนของความคลาดเคลื่อนสุ่มมีค่าไม่เท่ากัน โดยผู้ใช้สามารถระบุได้จากกล่องข้อความ Options ดังภาพที่ 4

ผลลัพธ์ในภาพที่ 5 เป็นการทดสอบค่าเฉลี่ยประชากรที่มีความแปรปรวนต่อค่าความคลาดเคลื่อนสุ่มมีความแปรปรวนไม่เท่ากัน ผลการทดสอบไม่ว่าจะพิจารณาจากสถิติบรวน์-ฟอร์ชิตหรือสถิติเวลช์ให้ค่า p-value น้อย

มาก และสรุปได้ว่าอุณหภูมิที่ต่างกันมีผลต่ออายุการใช้งานของยางสนที่ใช้หุ้มสายเชื่อมในวงจรเบ็ดเสร็จ

9. การใช้โปรแกรมสำเร็จรูป SAS

โปรแกรม SAS ส่วนที่เรียกว่า SAS Analyst [5] สามารถทำการทดสอบความแตกต่างระหว่างค่าเฉลี่ยประชากรในกรณีที่มีความคลาดเคลื่อนสุ่มมีความแปรปรวนต่างกันไปในแต่ละทรีทเมนต์ได้ โดยหลังจากป้อนข้อมูลแล้วให้เลือกเมนู



ภาพที่ 4 การเลือกสถิติบรวน์-ฟอร์ชิตและสถิติเวลช์จากกล่องข้อความ Options (โปรแกรม SPSS)

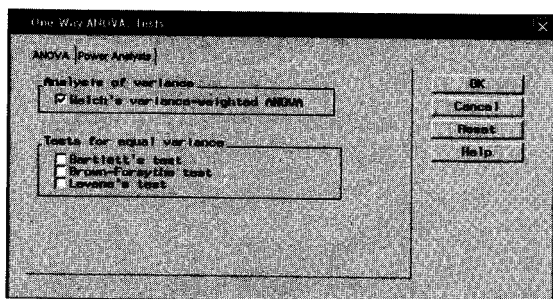
Robust Tests of Equality of Means

Y	Statistic ^a	df1	df2	Sig.
Welch	68.943	4	14.373	.000
Brown-Forsythe	111.745	4	18.298	.000

a. Asymptotically F distributed.

ภาพที่ 5 การทดสอบของบรวน์-ฟอร์ชิตและการทดสอบเวลช์

Statistics ตามด้วย ANOVA และเลือก One-Way ANOVA หลังจากนั้นให้ระบุตัวแปรตามและตัวแปรอิสระแล้วคลิกที่ปุ่ม Tests โปรแกรมจะแสดงกล่องข้อความ Tests ให้เลือกการทดสอบของเวลช์ (Welch's variance-weighted ANOVA) ดังแสดงในภาพที่ 6



ภาพที่ 6 การเลือกสถิติ Welch จากกล่องข้อความ Tests (โปรแกรม SAS)

The ANOVA Procedure

Welch's ANOVA for Y

Source	DF	F Value	Pr > F
TRT	4.0000	68.94	<.0001
Error	14.3730		

ภาพที่ 7 ผลลัพธ์เฉพาะการทดสอบของ Welch ในโปรแกรม SAS

สถิติ Welch และค่า p-value ที่ได้จากโปรแกรม SAS แสดงในภาพที่ 7 ซึ่งให้ค่าสถิติและ p-value ไม่ต่างจากโปรแกรม SPSS

10. บทสรุป

ทางเลือกอีกทางหนึ่งสำหรับการวิเคราะห์ข้อมูลในแผนแบบการทดลองสุ่มสมบูรณ์ในกรณีที่มีความแปรปรวนของความคลาดเคลื่อนสุ่มไม่เท่ากันในแต่ละทรีทเมนต์นั้น นอกจากใช้วิธีการแปลงข้อมูลให้มีความแปรปรวนของความคลาดเคลื่อนสุ่มเท่ากันในทุกกลุ่มแล้ว จากนั้นทำการวิเคราะห์ความแปรปรวนด้วยสถิติเอฟตามปกติ ซึ่งวิธีการนี้อาจก่อให้เกิดข้อสงสัยในเรื่องของการตีความหมายผลลัพธ์ที่ได้จากการทดสอบเนื่องจากผลการทดสอบของสถิติเอฟเป็นผลของการทดสอบบนมาตรวัดใหม่ไม่ใช่มาตรวัดเดิม หรืออาจใช้วิธีการกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักหรือใช้การทดสอบที่ไม่อิงพารามิเตอร์ก็ได้

อย่างไรก็ตาม ปัจจุบันโปรแกรมสำเร็จรูปทางสถิติได้เพิ่มสถิติที่ใช้ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากรซึ่งเป็นสถิติที่มีความแข็งแกร่งต่อการที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น โดยเฉพาะข้อที่ระบุว่าความแปรปรวนของความคลาดเคลื่อนสุ่มต้องมีค่าเท่ากันในทุกทรีทเมนต์ ซึ่งได้แก่ การทดสอบของบราวน์-ฟอร์ซิด และการทดสอบของเวลช์

อย่างไรก็ตามวิธีการทั้งสองนั้นไม่มีวิธีใดที่ให้ความถูกต้องแม่นยำสูงกว่าอีกวิธีหนึ่งโดยทั่วไป และไม่มีกฎเกณฑ์ตายตัวที่ระบุว่าควรใช้วิธีใดสำหรับขนาดตัวอย่างและความแปรปรวนค่าหนึ่งๆ ผู้ที่สนใจอาจศึกษาได้จากข้อเสนอแนะของ Cohen [6] ที่สรุปจากผลงานวิจัยของนักสถิติหลายท่าน เช่น โทมาร์กินและเซอร์ลิน (Tomarkin and Serlin ปี ค.ศ. 1986 อ้างถึงใน [6]) และ คลินช์และเคเซลแมน (Clinch and Keselman ในปี ค.ศ. 1982 อ้างถึงใน [6]) เป็นต้น

11. เอกสารอ้างอิง

- [1] Kuchl, R. O., Design of Experiments: Statistical Principles of Research Design and Analysis., 2 ed., Duxbury., Pacific Grove, 666 p, 2000.
- [2] Kutner, M. H., Nachtsheim, C. J., Neter, J., and Li, W., Applied Linear Statistical Models., 5 ed., McGraw Hill., Singapore, 666 p, 2005.
- [3] Oehlert, G. W., A First Course in Design and Analysis of Experiments., W.H. Freeman and Company., New York, 659 p, 2000.
- [4] Algorithmen fur SPSS 11.5. <http://www.rz.uni-hamburg.de/RRZ/Software/SPSS/Algorith.115/oneway.pdf> (September, 26, 2005)
- [5] SAS Institute Inc., The Analyst Application., SAS Institute Inc., Cary, 456 p, 1999.

- [6] Cohen, B. H., Explaining Psychological Statistics
(2nd Ed.), [http://www.psych.nyu.edu/cohen/](http://www.psych.nyu.edu/cohen/eps12dr1.pdf)
eps12dr1.pdf (September, 26, 2005)



สำนักหอสมุด